AN ANALYSIS OF RIBOSOME E-SITE BINDING PROTEINS AND THEIR FUNCTIONS IN TRANSLATION ELONGATION AND TERMINATION

by

Anthony P. Schuller

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, MD

October 2017

ABSTRACT

During translation, the ribosome uses substrate aminoacyl-tRNAs in concert with many protein factors to synthesize the protein encoded by a particular mRNA. Substrate tRNAs can bind inside the ribosome in three unique positions: A (aminoacyl), P (peptidyl), and E (exit). During peptide bond formation, the incoming aminoacyl-tRNA being decoded is accommodated into the ribosomal A site where it reacts with the P-site peptidyl-tRNA and is transferred to the A-site tRNA. Following a translocation step, the deacylated P-site tRNA moves to the ribosomal E site where it can be released back into the cytoplasm. Although we give these sites discrete names, communication between tRNAs and factors that bind in these sites has been of great interest to the field. Here I use a combination of *in vitro* and *in vivo* methods to understand the function of two E-site binding proteins in translation elongation and termination. First, using ribosome profiling and *in vitro* biochemistry I identify eIF5A, an abundant and essential protein in eukaryotes, as a global translation factor that stimulates both peptidyl-transfer (elongation) and peptidyl-hydrolysis (termination). Next, using an improved hydroxyl radical probing method I identify the ribosome binding site of Upf1, an essential factor of the nonsense-mediated mRNA decay pathway, as the L1 stalk near the ribosomal E site. Taken together, my observations expand on our understanding of communication between the ribosomal A, P, and E sites, as well as the impact of these factors on ribosome kinetics.

Rachel Green, PhD (Sponsor & Reader)

Professor

Department of Molecular Biology & Genetics

Johns Hopkins University School of Medicine

James Berger, PhD (Reader)

Professor

Department of Biophysics & Biophysical Chemistry

Johns Hopkins University School of Medicine

ACNKOWLEDGEMENTS

I have many people to thank for the incredible ride that this Ph.D. has been. First and foremost I would like to thank my advisor Rachel, whose constant support and curiosity has made me the scientist that I am. In particular, her commitment to the people she mentors has been truly inspiring and something I hope to emulate in my own career.

I would also like to thank the many mentors and collaborators that have contributed to my work and training throughout the years. These include my thesis committee comprised of James Berger (also reader), Geraldine Seydoux, and Jeff Corden who provided many great comments and experimental ideas throughout my Ph.D. Additionally, I had the opportunity to collaborate with many other wonderful scientists, including Roland Beckmann, Elena Conti, Bertrand Séraphin, Jun Liu, and the people in each of their labs.

I also have to thank the many members of the Green Lab, both present and past, and a few in particular: Laura Nevin for putting up with my incessant use of carts and constant expansion into her bench territory; Colin Wu for being my partner-in-crime as we investigated Upf1 and eIF5A; Karole D'Orazio for sharing her eIF5A prep during my hunt for contaminating proteins; Boris Zinshteyn for teaching me how to code; Kristin Smith-Koutmou for being my enzyme kinetics and biochemistry mentor; Eric Mills for being a great colleague, friend, and U of M fan; Aditya Radhakrishnan for being a great class- and lab-mate, and for introducing me to my wife Katie.

My parents have also been instrumental throughout my educational life. Their constant support and sacrifice throughout the years set me up to achieve my goals, and in so many ways I would not be the person I am without their love and support.

Last, but certainly not least, I need to thank my incredible, talented, brave, and compassionate wife (and partner-in-science) Katie. You have been an unending source of support and love, even when I told you I was going to “fractionate yeast lysate” as a fifth-year student. Your continual encouragement and optimistic attitude is inspiring, and I cannot wait to continue on our adventures in life and in science.

TABLE OF CONTENTS

LIST OF TABLES

Chapter II

Chapter III

LIST OF FIGURES

Chapter I

Chapter II

Chapter III

Appendix I

Chapter I: Introduction

The central dogma of biology refers to the flow of genetic information from DNA to RNA to protein. Across all domains of life, information conversion from RNA to protein (called translation) is performed by a large macromolecular machine called the ribosome. The ribosome, comprised mostly of RNA, reads the information encoded in an mRNA molecule, one codon at a time, and "translates" this information to synthesize the protein that the mRNA encodes. While on first glance, this process may seem rather simple; however, it requires the intricate interplay of many molecules including the ribosome, mRNAs, tRNAs, and other protein factors that all come together with precise timing to achieve regulated gene expression.

The process of translation can be broken into four main phases: initiation, elongation, termination, and recycling (Figure 1). Translation is highly conserved between eukaryotes, bacteria, and archaea, although there are substantive differences as well. In this chapter, I will focus on eukaryotic translation, and in particular elongation, termination, and recycling. During initiation, many protein initiation factors (eIFs) guide the proper assembly of an 80S ribosome positioned at the AUG start codon with an initiator methionyl-tRNA bound in the P site. During elongation, this 80S ribosome moves processively along the mRNA, three nucleotides per step, synthesizing the encoded protein one amino acid at a time through the coordinated action of aminoacyl-tRNAs and elongation factors (eEFs). At the end of the protein-coding region (also called an open reading frame, or ORF) the ribosome encounters a termination codon, which is specifically recognized by a set of protein factors called release factors (eRFs) that

promote the release of the nascent protein from the peptidyl-tRNA (and ultimately the ribosome). Finally, in the recycling phase, the terminated 80S ribosome is dissociated into separate 40S and 60S subunits to begin a new round of translation.

In this chapter, I cover recent advances in our understanding of translation elongation, termination, and recycling. I outline our current understanding of the order and timing of events that occur at each phase of translation, paying particular attention to recently identified events or regulation that impact each phase, such as ribosome stalling during elongation, the role of mRNA-context in termination, and ribosome rescue. At the heart of each step is a kinetic decision that the ribosome must make: to translate or not to translate. In some cases, there are auxiliary factors that resolve pauses, such as eIF5A, while in other instances the ribosome may abandon the translation cycle altogether with the help of ribosome rescue factors Pelota:Hbs1L (yeast Dom34:Hbs1). These decisive moments can have a profound impact on the output from a particular mRNA, determining how much or which products might be produced. As with all processes, kinetics and thermodynamics driven by cellular concentrations of players in the process dictate biological outcome.

Elongation

The process of translation elongation begins immediately after initiation has taken place and an 80S ribosome is positioned at an AUG start codon with a methionyl-$tRNA^{iMet}$ in the P site (Figure 2A). While I will not cover initiation in this chapter, a number of excellent reviews already exist on this topic (Hinnebusch, 2014; Hinnebusch and Lorsch, 2012; Jackson et al., 2010; Sonenberg and Hinnebusch, 2009). The eukaryotic elongation phase for the ribosome extends from the loading of the first aminoacyl-tRNA at the start of the ORF (after the initiation codon) until the ribosome reaches the termination codon at the end of the ORF, and is thought to be mostly conserved relative to bacterial elongation, as outlined below (Dever and Green, 2012; Voorhees and Ramakrishnan, 2013).

Elongation begins with the tRNA selection process wherein the aminoacyl-tRNA that contains the proper anticodon to match the mRNA codon is loaded into the A site. Aminoacyl-tRNAs are delivered to the ribosomal A site by a specialized GTPase called eEF1A (EFTu in bacteria) in a ternary complex with GTP (Carvalho et al., 1984; Fischer et al., 2015; Pape et al., 1998; Shao et al., 2016). Once cognate interactions between the codon and anticodon are sensed, GTPase activation and hydrolysis take place, eEF1A:GDP is released from the ribosome, and the cognate tRNA is fully accommodated into the A site. From this position, the amine moiety of the amino acid on the aminoacyl-tRNA nucleophilically attacks the aminoacyl ester linkage on the P (peptidyl)-site tRNA and the growing peptide chain is transferred to the aminoacyl-tRNA. As peptide bond formation occurs, the ribosomal subunits are thought to undergo rotation with respect to one another and the tRNAs adopt an altered conformation referred to as the hybrid states

of binding (P/E and A/P, respectively) – in this state, the anticodon end of the tRNAs remain positioned essentially in the P and A sites of the small subunit, while the acceptor ends of the tRNA are positioned in the E and P sites of the large subunit (Moazed and Noller, 1989). This rotated state of the ribosome becomes the substrate for the action of another specialized GTPase (eEF2; EFG in bacteria), which translocates the mRNA:tRNA complex relative to the ribosome, and returns the tRNAs and ribosome to their so called classical states (E/E and P/P, and un-rotated, respectively) (Ferguson et al., 2015; Rodnina et al., 1997; Sengupta et al., 2008; Taylor et al., 2007). In addition to eEF1A and eEF2, an additional factor in fungi (called eEF3) has been shown to be essential for proper elongation, potentially by promoting tRNA release from the E site after translocation (Andersen et al., 2006; Triana-Alonso et al., 1995). This cycle is repeated over and over again until each codon has been translated and the full protein synthesized.

As the ribosome elongates along an ORF, it can encounter a variety of problematic sequences that can negatively affect the process. First, certain amino acid combinations have been shown to stall the ribosome, either because of poor reaction kinetics or because the nascent polypeptide chain adopts inhibitory conformations in the exit tunnel. Other problems that the ribosome encounters include codon-based stalling events, which most typically occur when an mRNA includes a codon that is "rare," and thus whose tRNA is underrepresented in the pool of available tRNAs. In other quite particular cases, the order in which certain codons appear is critical for stalling, suggestive of complexity in the interactions of certain tRNAs within the ribosome. Additionally, the ribosome may encounter mRNA secondary structure motifs (such as

stem loops or pseudoknots) or "slippery sequences" that can cause elongation arrest and frameshifting. In all of these cases, the ribosome has a decision to make: to continue or abort translation. Several of these challenging events are corrected by simply waiting – the ribosome stalls and waits until the proper factor or tRNA is delivered. In other cases, the ribosome undergoes a frameshifting event in order to resume translation, now in a different frame where the problematic sequence is no longer relevant. Below I review our current understanding of various elongation problems and how the ribosome or other factors are able to resolve these kinetic challenges and redirect the ribosome to ongoing translation.

Amino acid stalling

During translation elongation, the ribosome is faced with the challenge of making 400 unique dipeptide products within the same active site. This number comes from the combination of twenty different amino acids that can be found on the P- and A-site tRNAs. Unlike many molecular machines that are specific for one substrate or reactant, the ribosome must be flexible enough to allow for reactivity between these twenty substrates and twenty reactants in the face of peptidyl-tRNA structure that may result from higher order peptide structure. Over the course of decades of molecular biology and biochemical work, it has been established that while the ribosome is capable of making all 400 possible dipeptide linkages, not all reactions are equally favorable (Figure 2B). For example, proline is a unique "imino" acid where the reactive amine (of the amino acid) is found within a five-membered nitrogen-containing ring, and thus its nucleophilicity (as an acceptor in the A site) is substantially reduced (Pavlov et al., 2009;

Wohlgemuth et al., 2008). Additionally, because of entropic constraints, proline may not be an ideal donor substrate in the P site. As a result of these different negative contributions to catalysis, translation of Pro-Pro bonds is slow, and the addition of a third proline is even more dramatically challenging for the ribosome (Doerfel et al., 2013; Gutierrez et al., 2013; Ude et al., 2013). In addition to poor reactive chemistry in the actual peptidyl-transferase center, poly-proline structure inside the ribosome exit tunnel may make inhibitory interactions with tunnel residues, as has been identified for other ribosome stall sequences (Wilson et al., 2016).

While poly-proline formation is kinetically slow, stretches of proline are found throughout eukaryotic genomes, and the encoded proteins are indeed translated. Initial work in *E. coli* identified a ribosome-interacting protein named EFP to be essential for resolving translation stalls at proline stretches (Doerfel et al., 2013; Glick et al., 1979; Glick and Ganoza, 1975; Ude et al., 2013). Shortly after this discovery, it was identified that eIF5A, the eukaryotic homolog of EFP, was also essential for the translation of poly-proline sequences (Gutierrez et al., 2013). eIF5A is a small, highly abundant, and essential protein in eukaryotes, comprised of only 157 amino acids and containing a unique post-translational modification called hypusine (Park et al., 1981). eIF5A is a historical name given to this protein which was identified through fractionation of yeast lysate and an associated biochemical activity in stimulating a Met-Puromycin reaction (Benne and Hershey, 1978; Kemper et al., 1976; Schreier et al., 1977). In retrospect, this biochemical assay is more simply reflective of peptide bond formation than true initiation, and work from multiple groups over the past few decades has suggested that eIF5A's predominant role in translation is in generally promoting elongation (Gregio et

al., 2009; Gutierrez et al., 2013; Henderson and Hershey, 2011; Saini et al., 2009). In any case, eIF5A was shown to be critical for the ribosome to overcome the kinetic defect of Pro-Pro bond formation, dependent on its unique hypusine modification (REF Gutierrez). Indeed, recent structural work suggests that this modification specifically helps to stabilize the conformation of the peptidyl-tRNA for nucleophilic attack by the A-site aminoacyl-tRNA (Melnikov et al., 2016; Schmidt et al., 2016).

While stretches of three or more consecutive prolines are found in as few as 10% of yeast genes, it was initially argued that the essential nature of eIF5A derives from this specific function in translation (Gutierrez et al., 2013). However, more recent unbiased approaches to define the *in vivo* function of eIF5A by ribosome profiling revealed significant pausing across a wide spectrum of amino acid motifs, including in particular those containing proline, aspartic acid, glycine, alanine, valine, and isoleucine residues (Pelechano and Alepuz, 2017; Schuller et al., 2017). Experiments using an *in vitro* reconstituted translation system verified that eIF5A is critical for the translation of all sequences tested, including those containing the most problematic and the least problematic amino acids (Schuller et al., 2017). These data together suggested that eIF5A is a universal translation factor that functions during the formation of each peptide bond, increasing the processivity and efficiency of the overall process (Figure 2A-B).

These findings also suggest a critical role for the ribosome E site as a sensor of ribosome kinetics. After peptide bond formation and translocation, the deacylated tRNA is shifted to the ribosome E site, and the A site is then open for decoding. Following accommodation of the next aminoacyl-tRNA, the deacylated tRNA is released from the E site (Gnirke et al., 1989). In the case of slow peptide bond formation (such as Pro-Pro),

this aminoacyl-tRNA would remain unreacted, but the E site would be available for eIF5A binding to promote peptide bond formation. In this way, the unoccupied E site is a sensor of translation arrest and eIF5A binding at this site functions to resolve this kinetic stall.

Codon-based stalling

In addition to particular amino acids causing translation elongation problems, ribosome stalling can also occur when the ribosome encounters specific "rare" or "suboptimal" codons in the open-reading frame (Figure 2C) (Plotkin and Kudla, 2011; Quax et al., 2015). The idea of "rare" codons has been discussed in the translation field for many years (Ikemura and Ozeki, 1983), and it has been observed that organisms preferentially use certain codons relative to others to encode a particular amino acid (Ikemura and Ozeki, 1983; Plotkin and Kudla, 2011; Quax et al., 2015). A particular codon's "optimality" is essentially a reflection of that codon's usage in the transcriptome and the availability of the corresponding tRNA for use by translating ribosomes (dos Reis et al., 2004; Pechmann and Frydman, 2013; Sharp and Li, 1987). Codons that are "optimal" on this scale have a readily available pool of tRNAs for translation elongation; conversely, "suboptimal" or "rare" codons have a limited supply of the corresponding tRNA for translation. Translation stalls at these codons can result from two distinct routes: 1) low cognate tRNA levels (Dana and Tuller, 2014; Gardin et al., 2014; Ikemura and Ozeki, 1983) or 2) poor aminoacylation by the corresponding aminoacyl-tRNA synthetase and therefore a reduction in the cognate aminoacyl-tRNA level (Elf et al., 2003). In both cases the ribosome attempts to decode the A site codon but does not

receive the cognate aminoacyl-tRNA, thereby creating an open A site (and translation arrest).

Over the past decade, several groups have identified a correlation between the collective "optimality" of an mRNA (e.g. the average optimality of codons over the entire ORF) and translation rate (Dana and Tuller, 2014; Gardin et al., 2014; Yu et al., 2015). Other work has identified correlations between optimality and mRNA stability (a more in-depth discussion of this can be found elsewhere (Hanson and Coller, 2017)). The impact of a "suboptimal" codon derives from a kinetic barrier that the ribosome faces when the limiting tRNA leaves the A site unoccupied. At this delay, the ribosome may stall until the proper tRNA is delivered to the A site, or until a near- or non-cognate tRNA is delivered and a non-canonical event occurs in order to proceed (such as frameshifting) (Gallant and Lindsley, 1998; Gurvich et al., 2005; Kane, 1995; Temperley et al., 2010), or the ribosome may be targeted for downstream quality control (Graille and Seraphin, 2012) .

Distinct from "rare" codon-based stalling, translation elongation can also be impacted by the particular order of codons, even when one or both of them is not "rare." A recent study using FACS analysis with a randomized GFP library containing three adjacent, random codons identified 17 different codon pairs that inhibited translation (Gamble et al., 2016). In some cases, this inhibition could be explained by the low abundance of a particular tRNA or by inefficient wobble decoding by the tRNA. In other instances, the authors observed that the particular order of a codon pair impacts whether it permits or precludes optimal translation (Figure 2D). For a set of twelve codon pairs, the GFP output of each pair was compared to an optimal pair of the same amino acid

composition, and it was discovered that certain codon orders are inhibitory to translation compared to their reverse order counterparts. For example, the pair CUC-CCG (Leu-Pro) was strongly inhibitory compared to the optimal, non-inhibitory UUG-CCA (Leu-Pro). Moreover, in the reverse order, CCG-CUC (Pro-Leu) is not inhibitory, but just as "optimal" as the control CCA-UUG (Pro-Leu). These observations are suggestive of subtle coordination between the P and A site tRNAs during elongation, and highlight the incredible complexity of the genetic code. Further investigation into the mechanism of this communication and how the ribosome or other translational machinery is critical for resolving this type of elongation stall will be an interesting source of continued exploration.

mRNA structure-based stalling

Another problematic circumstance the ribosome may face during translation elongation is the presence of mRNA secondary structure in an mRNA. The presence of an RNA stem loop or pseudoknot structure has been shown to cause translation stalling. In a subset of these circumstances, structure in the mRNA has evolved such that the ribosome is found at a "slippery sequence," such as AAAAAAG, that in turn can promote ribosome frameshifting (Belew et al., 2014; Biswas et al., 2004; Caliskan et al., 2017; Namy et al., 2006; Yan et al., 2015). These endogenous events that are clearly beneficial to the system are referred to as programmed ribosomal frameshifting (PRF) (Dinman, 2012). While the ribosome stall in PRF is similar to that of other elongation stalls discussed earlier, this circumstance is unique in that it is advantageous for the desired gene expression outcome.

Viruses are common users of PRF, allowing them to more efficiently pack information into their genome as their size is often restricted by capsid size. Often, viruses use PRF to control the ratio of structural and enzymatic proteins being translated, such as the case of HIV-1 where a -1 PRF helps to control the relative ratios of Gag and Gag-Pol synthesis for proper viral particle replication and assembly (Jacks et al., 1988). The HIV-1 PRF signal is comprised of two parts: a downstream mRNA stem loop structure and a "slippery site" that is U-rich (Biswas et al., 2004). The mRNA structure leads to ribosome stalling over the slippery site, the slowed kinetics increasing the likelihood of a frameshift event that permits translation of the downstream, out-of-frame pol gene. Yet again the ribosome is faced with a kinetic difficulty (in this case a large RNA stem-loop), but in a fashion distinct from other stalling events, PRF leads to a ribosome frameshift that can get past the kinetic barrier and continue translation of the downstream gene.

Termination

The process of translation termination begins when the ribosome encounters a termination codon in the ribosomal A site (Figure 3A) (Dever and Green, 2012). In eukaryotes, termination is encoded by three codons: UAA, UAG, and UGA, all of which are recognized by the eukaryotic release factor eRF1. The overall shape and size of eRF1 is strikingly similar to that of a tRNA (expected of something that binds the same site) with two distinct functional ends. First, like tRNA, eRF1 contains a structural motif that recognizes with high specificity all three termination codons (called NIKS motif) (Brown et al., 2015; des Georges et al., 2014; Frolova et al., 1994; Matheisl et al., 2015; Song et al., 2000). Second, like tRNA, eRF1 has precisely positioned GGQ motif that extends into the peptidyl transferase center to promote catalysis. And, finally, like tRNA, eRF1 is delivered to the ribosomal A site by a specialized GTPase called eRF3 (Frolova et al., 1996) that is related to EFTu. While structurally distinct from eEF1A (Kong et al., 2004), eRF3 is required for multiple turnover reactivity of eRF1, and it delivers eRF1 to the ribosome in a GTPase-dependent manner (Alkalaeva et al., 2006; Eyler et al., 2013; Salas-Marco and Bedwell, 2004). Once eRF3 delivers eRF1 to the ribosomal A site, the eRF1 GGQ motif coordinates a water molecule at the ribosome peptidyl-transferase center that will hydrolyze the P-site peptidyl-tRNA (Frolova et al., 1999). Upon release of the peptide, the process of translation termination is complete.

Recent structural work has revealed that eRF1 not only interacts with the three nucleotides of the termination codon, but also with the +4 nucleotide of this sequence (Brown et al., 2015; Matheisl et al., 2015); these structural observations are nicely consistent with the observation from ribosome profiling that the ribosome protected

fragment (RPF) size is one nucleotide longer at stop codons (Ingolia et al., 2011). The dependence of termination specificity on a fourth nucleotide was established years ago using computational methods (Brown et al., 1990) and more recently using reporter mRNAs in yeast (Bonetti et al., 1995) and mammalian cells (Floquet et al., 2012). These studies revealed that particular termination codons (and +4 nucleotide identities) were more or less likely to promote translation read-through, suggesting a more complicated code for termination than the simple three-nucleotide codon. "Weaker" termination codons (such as UGA/C) caused increased ribosome read-through events, presumably by incorporation of near-cognate aminoacyl-tRNA or by frameshifting, compared to a "stronger" termination codon (UAA/G). Taken together, while all three stop codons can elicit translation termination, the +4 nucleotide position creates an added level of termination regulation in the mRNA coding sequence and could have effects on overall gene expression regulation or lead to unique C-terminal protein extensions (Arribere et al., 2016; Schueren and Thoms, 2016).

While there are many broad similarities between translation termination and elongation, an additional similarity comes from the influence of eIF5A. From a biochemical mechanism standpoint, translation termination and elongation are quite similar: in the case of peptidyl-transfer, a nucleophilic attack occurs by the amino acid conjugated on the incoming A-site tRNA while in termination, nucleophilic attack occurs by the water molecule coordinated by the GGQ motif eRF1 (Figure 3B). From a kinetic standpoint, it has been shown that the rate of peptidyl-release is much slower than the rate of peptidyl-transfer (Youngman et al., 2004), potentially because the water molecule is a worse nucleophile than most amino acids. In a recent study, I found that eIF5A strongly

enhanced the rate of both peptidyl-transfer and peptidyl-release (Schuller et al., 2017). As the chemistry of these two events is quite similar, it is not surprising that this is the case; however it does highlight the fact that translation termination is kinetically slow compared to elongation and further suggests a role for the E site as a sensor of ribosome kinetics. In this case, termination is kinetically slow, the E site is unoccupied, and eIF5A can bind to stimulate peptidyl-release. Taken together with its great abundance (>273,000 copies per cell in yeast; (Kulak et al., 2014)) and high affinity for ribosomes (Rossi et al., 2016), eIF5A is likely present on most (if not all) ribosomes throughout translation.

mRNA-context controls termination

While the specifics of eRF1:eRF3 binding at the ribosomal A site, and the mechanism of peptidyl-release have been well determined, recent work has begun to shed light on more subtle specification of the termination process. As previously mentioned, the +4 nucleotide interactions between eRF1 and the mRNA termination codon suggest added gene expression complexity at the mRNA coding sequence. If certain stop codons allow for frameshifting or mis-incorporation more frequently, you could imagine that this could produce alternative C-termini on proteins which could in turn affect functionality or stability (Arribere et al., 2016; Schueren and Thoms, 2016). At the heart of such events, as with the others I have reviewed, is a kinetic delay that the ribosome encounters. For instance, a weak termination codon (UGA/C for example) could recruit eRF1 less efficiently, causing the ribosome to spend more time with an empty A site. At this point the ribosome has two options: 1) to wait for eRF1:eRF3 or 2) to undergo frameshifting or mis-incorporation events to continue translating (Roy et al., 2015). This decision could be

influenced by other factors, such as eRF1 concentrations or GTP availability for eRF3, but in either case has profound repercussions for the protein being translated.

While this first example relies on direct eRF1:ribosome interactions, it is also possible that other proteins, including mRNA binding proteins, could affect the termination process through interactions with release factors or the ribosome. Recent work on the poly(A)-binding protein (PABP) has suggested a potential influence on the efficiency of termination. Biochemical interactions between PABP and the N-terminus of eRF3 have been documented in the literature for many years (Hoshino et al., 1999; Ivanov et al., 2008) though the consequence of this interaction was unclear. Using an *in vitro* reconstituted termination system, it was recently shown that PABP can directly promote the recruitment of eRF1:eRF3 to a terminating ribosome (Ivanov et al., 2016). One hypothesized consequence of the PABP:eRF3 interaction was that it could lead to an increase in localized eRF1:eRF3 concentration near the terminating ribosome (and therefore more efficient termination). This study, however, found that PABP promoted recruitment of eRF1:eRF3 to the ribosome independent of mRNA-binding activity, suggesting PABP may influence the conformation of eRF1:eRF3 on the ribosome. Indeed, other *in vitro* studies have shown that stop codon recognition is a two-step process, potentially including a conformational change in the ribosome (Alkalaeva et al., 2006; Kryuchkova et al., 2013).

Additional connections between PABP and termination come from investigations of the nonsense-mediated decay (NMD) pathway in eukaryotes. NMD is a cellular quality control pathway that that contain what is broadly referred to as a premature termination codon (PTC) dependent on the three conserved Upf proteins: Upf1, Upf2,

and Upf3 (Kervestin and Jacobson, 2012; Popp and Maquat, 2013). These mRNAs are created by multiple routes: genes may carry a mutation that results in a PTC (Holbrook et al., 2004), inefficient splicing may lead to export of a pre-mRNA with a PTC (almost inevitably) encoded in the intron (He et al., 1993; Mitrovich and Anderson, 2000), or the stop codons of upstream open reading frames (Mendell et al., 2004; Welch and Jacobson, 1999) and non-coding RNAs (Marquardt et al., 2011; Smith et al., 2014) can be sensed as PTCs.

In many eukaryotes, the presence of a termination codon upstream of an exon-junction complex (EJC; a large protein complex deposited at splice junctions) is known to be a strong signal for NMD (Le Hir et al., 2001). Upf2 has been shown to interact with components of the EJC and Upf1 (Le Hir et al., 2001; Melero et al., 2012), potentially recruiting Upf1 to a nearby terminating ribosome to signal NMD. However, some organisms that lack an EJC (such as *S. cerevisiae*) have robust NMD, suggesting a more common mechanism than direct recruitment by the EJC.

This common thread may be mediated by poly(A)-binding protein (PABP) as it has been shown to affect NMD efficiency (Amrani et al., 2004). As PABP has been shown to stimulate translation termination directly (Ivanov et al., 2016), it is possible that the distance between the PTC and PABP can determine whether the mRNA undergoes NMD or not. In fact, PTCs located near the 3'end of open-reading frames have been shown to be weaker substrates for NMD (ie. the mRNAs are more-stable).

Taken together these ideas suggest a potential common mechanism between context-based termination and NMD (Figure 4). If a termination codon is located sufficiently close to PABP (or other 3'UTR binding proteins), the ribosome would

terminate efficiently, and the mRNA remain stable. If a termination codon were premature (and far from PABP), termination would be inefficient, leading to ribosome stalling and recruitment of NMD machinery potentially influenced also by the presence of an EJC.

The idea that mRNA context (both sequence and protein-context) can impact translation termination has recently come into the limelight with the discovery of ciliate and trypanosomal organisms that use the canonical termination codons (all of them) to code for amino acids at certain sites and as true termination codons at other sites (Heaphy et al., 2016; Lobanov et al., 2017; Swart et al., 2016; Zahonova et al., 2016). For some of these organisms, transcriptome sequencing revealed that the canonical termination codons could be simply used to encode amino acids as there are "cognate" tRNAs for the termination codons encoded in their genome. Further observations by bioinformatics and the use of ribosome profiling created a model where the termination codon (UAA, UAG, UGA) is used to encode an amino acid unless it is located sufficiently close to the mRNA poly(A) tail (and presumably PABP) (Figure 3C) (Heaphy et al., 2016; Lobanov et al., 2017; Swart et al., 2016). The strongest support for model comes from the fact that many mRNAs in these organisms contain short (or nonexistent) 3'UTRs, however a direct connection to PABP remains to be shown.

Recycling

At the end of termination, the 80S ribosome (now containing only a deacylated-tRNA in the P site) must be recycled into 40S and 60S subunits before beginning the next round of initiation (Figure 5A). Subunit recycling is accomplished by the ribosome-recycling factor ABCE1 (ATP-binding cassette protein E1; also called Rli1 in *S. cerevisiae*). ABCE1 is an essential protein in all eukarya, containing two nucleotide-binding domains (NBDs) and an N-terminal iron-sulfur (FeS) cluster (Karcher et al., 2008). Using the force of ATP-binding and hydrolysis, ABCE1 dissociates the post-termination ribosome into 40S and 60S subunits with the coordination of release factor eRF1 (Barthelme et al., 2011; Pisarev et al., 2010; Shoemaker and Green, 2011). In *S. cerevisiae*, there is also data that suggests ABCE1:eRF1 interactions couple translation termination and recycling, as ABCE1 was found to stimulate the actual catalytic activity of eRF1 (Shoemaker and Green, 2011).

Ribosome profiling experiments in an ABCE1-depletion strain have confirmed this role in subunit recycling *in vivo*, and highlight the consequence of improper recycling: translation reinitiation in the 3'UTR (Young et al., 2015). Upon ABCE1 depletion, increased ribosome occupancy was observed both at the termination codon (suggestive of a defect in termination/recycling) and downstream in the 3'UTR. Additional experiments using amino acid starvation and reporter constructs revealed that these downstream ribosomes were indeed translating, having reinitiated near the main ORF stop codon, sometimes in a different frame. These data suggested that unrecycled 80S ribosomes act similarly to those stalled in elongation with an empty A site (Figure

2C). In both cases, the ribosome resolves this problem by frameshifting to return to elongation.

While the role of ABCE1 in recycling is clearly critical to the translation process, before ABCE1 was discovered to function in recycling it was implicated in translation initiation. Work from several labs identified interactions between ABCE1 and initiation factors including eIF2α, eIF2, eIF3, and eIF5 (Andersen and Leevers, 2007; Chen et al., 2006; Dong et al., 2004). Additionally, it was observed that ABCE1 sediments with 40S subunits in a sucrose gradient dependent on the ATPase activity (Andersen and Leevers, 2007; Dong et al., 2004). Finally, in yeast, genetic depletion of ABCE1 was found using polysome analysis to cause impaired assembly of ribosome pre-initiation complexes, and led to overall reduced rates of bulk translation initiation as observed by a reduction in polysomes (Dong et al., 2004). As the process of ribosome recycling happens immediately upstream of initiation in the translation cycle, it has been plausible that ABCE1 plays important functional roles in both events.

Recent structural data for ABCE1 bound to 80S ribosomes (Brown et al., 2015; Preis et al., 2014) and 40S subunits (Heuer et al., 2017) has provided the field with a detailed view of how ABCE1 may function to stimulate ribosome recycling and recruit translation initiation machinery for the next round of translation. In the 80S:ABCE1 pre-splitting structure, the ABCE1 FeS cluster is found positioned in the ribosomal A site, making direct contacts with the eRF1 C-terminal domain (Brown et al., 2015; Preis et al., 2014). In a post-splitting 40S:ABCE1 structure, the FeS cluster adopts a completely different conformation, rotated approximately 150 degrees from the pre-splitting state (Heuer et al., 2017). Superposition of these structures suggests that, during splitting, the

FeS cluster movement (likely coordinated by ATP binding or hydrolysis) forces eRF1 deeper into the inter-subunit space of the ribosome, thereby causing dissociation of the subunits. After splitting, the position assumed by the FeS cluster would clash with uL14 of the 60S subunit, thereby preventing 60S rejoining and ensuring proper recycling.

A 40S:ABCE1 state was also observed in a different study that presented the structure of a 48S pre-initiation complex from mammalian cells (Simonetti et al., 2016). While the ABCE1 density in this structure was originally attributed to eIF3i and g, it has since been corrected as ABCE1 (Mancera-Martinez et al., 2017). As this structure includes other members of the pre-initiation complex, including subunits of eIF2 and eIF3, it seems likely that ABCE1 might play a dual-role in promoting ribosome recycling and re-initiation through initial interactions with the 80S ribosome and downstream more specific interactions with the 40S subunit. Further mechanistic and structural studies will be required to fully understand these connections.

Ribosome Rescue

As the ribosome faces problematic encounters throughout translation, it can resolve these issues in a productive manner (as discussed throughout this chapter), or a process known as ribosome rescue may be required as a last resort. In eukaryotes, there is a complex set of systems that target the improper proteins resulting from translation arrest, degrade the problematic mRNAs, and rescue ribosomes (Graille and Seraphin, 2012; Nurenberg and Tampe, 2013; Shoemaker and Green, 2012). Here I will focus on ribosome rescue and its mechanism, much of which is shared with ribosome recycling.

Early genetic studies in yeast identified the gene DOM34 as critical for selective degradation of a stemloop-containing mRNA in the No-Go decay (NGD) pathway (Doma and Parker, 2006). While I will not discuss NGD in depth, NGD is a broadly defined mRNA quality control pathway that degrades mRNAs with sequence features that inhibit translation (Shoemaker and Green, 2012). Biochemical work using an *in vitro* reconstituted translation system identified that Dom34 (and associated GTPase Hbs1) functioned to dissociate ribosome subunits independent of peptidyl-release (Shoemaker et al., 2010). Dom34 (Pelota in mammals) is structurally homologous to eRF1 (Graille et al., 2008; Lee et al., 2007) with three important exceptions: 1) it does not contain the NIKS motif for stop codon recognition, 2) it does not contain the GGQ motif for peptidyl-hydrolysis, and 3) it is delivered to the ribosome by a distinct GTPase called Hbs1 (Carr-Schmid et al., 2002; Chen et al., 2010). These structural differences create a unique role for Pelota, allowing it to bind in the A site of ribosomes in a codon independent manner and rescue ribosomes without releasing the peptide chain (as eRF1 does).

In the cell, ribosome profiling experiments using a Dom34 knockout strain revealed that Dom34 functions to rescue ribosomes found in the 3'UTR or the poly(A) tail, or at the end of truncated mRNAs resulting from endonucleolytic cleavage (Figure 5B) (Guydosh and Green, 2014, 2017). Additional experiments combining ABCE1 depletion and Dom34 overexpression further elucidated this connection between impaired ribosome recycling and Dom34-mediated ribosome rescue (Young et al., 2015). This functional evidence is not limited to ribosome profiling, but also is supported by genetic studies using truncated mRNA reporter constructs that connect Dom34:Hbs1 to

the NGD pathway that recognizes ribosomes stalled at the 3' ends of truncated mRNAs (Tsuboi et al., 2012). Taken together these observations suggest a general role for Dom34 to rescue or "clean-up" ribosomes when they translate into regions they should not, including 3'UTRs, and poly(A) tails, or when the ribosome simply cannot translate any further on a truncated message.

Though mechanistically different, Dom34 and eRF1 both interact with ABCE1 in order to achieve full kinetic activity. Although Dom34 is sufficient to promote subunit dissociation *in vitro* (Shoemaker et al., 2010), it was observed that ABCE1 greatly enhanced the rate of subunit recycling (Pisareva et al., 2011; Shoemaker and Green, 2011) using the same force of ATP-binding and hydrolysis as in coordination with eRF1. This coordination was also observed in *S. cerevisiae* (Rli1:Dom34) where these proteins were shown to recycle inactive 80S ribosomes bound by Stm1 (Figure 5C) (van den Elzen et al., 2014). Stm1 is a small protein that binds in the mRNA channel at the 80S subunit interface upon glucose deprivation, and causes the formation of vacant, inactive 80S ribosomes. These inactive ribosomes cause a decrease in the cytoplasmic pool of available ribosomes for translation initiation, and lead to the specific translation of stress response genes that assist the cell during this stress. Mechanistically, Stm1-bound ribosomes are very similar to those found at the end of a truncated mRNA, with an open A site available for Pelota to bind.

Although this function for Dom34 (Pelota in mammals) in rescuing aberrant translation products or stalls may seem simply as a "clean-up" mechanism for the cell, this ABCE1:Pelota activity has also been shown to be critical for ribosome homeostasis and translational output in certain blood cell lineages. A recent study discovered that

during K562 differentiation (upon addition of hemin), the levels of ABCE1 decrease, concomitant with an increase in 3'UTR ribosomes which are not properly recycled (Figure 6) (Mills et al., 2016). During the initial phase of ABCE1 loss as the cells differentiate toward erythrocytes, the cells quickly upregulate the ribosome rescue factor Pelota which allows for increased recycling of the 3'UTR ribosomes, and translational reprogramming to synthesize hemoglobin. However, as erythrocytes are enucleate cells, they are not able to produce more ribosomes or rescue factors to continuously account for this loss of recycling activity. This creates a situation where Pelota:Hbs1L levels decrease over time from standard protein degradation pathways, and 3'UTR ribosomes are again accumulated. Without ABCE1 or Pelota to recycle ribosomes, this 3'UTR accumulation eventually leads to a global translation defect and a trend toward decreased hemoglobin output after 10 days (on the order of the life-span of a red blood cell) (Shemin and Rittenberg, 1946).

This study further sought to understand the effect of ribosome rescue factors when there is a cellular perturbation on the available ribosome pool. "Ribosomopathies" are a heterogeneous set of diseases that result from perturbations in ribosome homeostasis, such as the loss of a ribosomal protein, that can present as hematopoietic dysfunction (Narla and Ebert, 2010). Diamond-Blackfan Anemia is a well-documented "ribosomopathy" that results from the heterozygous loss of ribosomal protein rpS19, and is associated with a reduced erythroid progenitor cell population (Draptchinskaia et al., 1999). Mills et al. tested the connection between ribosome rescue factors and the loss of rpS19, finding that overexpression of Pelota:Hbs1L was required to rescue the global translational defect caused by rpS19 depletion (and associated depletion of the total

cellular ribosome pool). These results suggest that a critical regulation of ribosome rescue factors to rescue 3'UTR ribosomes (resulting from improper or overwhelmed rescue machinery) is important to maintain a well-equipped cytoplasmic pool of active ribosomes for translation homeostasis.

CONCLUSION

While the ribosome was discovered more than sixty years ago and the genetic code completed in 1966, the field continues to unravel the complexity of translation and gene expression regulation that exists beyond the simple three-letter codon. As the ribosome proceeds through the phases of translation (Figure 1), it is forced to make many decisions, each determined by many factors, but what unites these influences is their impact on kinetics.

During elongation (Figure 2), the ribosome must work to synthesize the peptide encoded by the mRNA codons, even when particular sequences of amino acids, codons, or interactions between tRNAs impose kinetic barriers. In some instances, the cell has evolved intricate machinery to promote faster translation kinetics (eIF5A, for example), while in others, the cellular mechanism to resolve these kinetic defects remains unknown. For instance, the discovery of specific tRNA/codon pairs that inhibit translation elongation remains quite puzzling. What is it about these particular pairs elicits a translation arrest by the ribosome? And why does the order of the tRNA/codon pair matter? Moving forward, it will be important to better understand the communication between the ribosomal P and A sites during elongation, and how both codons and tRNAs between sites communicate with the ribosome. Additionally, the role of the ribosome E site remains to be fully defined. The global stimulatory factor eIF5A must compete with E-site tRNA for binding and stimulation of peptide bond formation. In this way, the occupancy of the E site could be an indicator of the translation elongation rate at a particular site in the coding sequence. Deeper investigation into the coordination of the E, P, and A sites using careful biochemical or single-molecule approaches should provide a

deeper understanding of the dynamics during these elongation stalls, and their influence on gene expression regulation.

Translation termination (Figure 3) is also a rich area for continued exploration. What are the determinants of stop codon recognition beyond a single codon, for example, as specified by local mRNP (messenger ribonucleoprotein) composition? Recent studies in organisms that use termination codons for both coding and termination functions (sense and nonsense) has led to increased interest in models where termination is guided by the mRNP context near the termination codon. One potential candidate that may directly stimulate termination is the poly(A)-binding protein (PABP) (Ivanov et al., 2016). While both computational and biochemical approaches have begun to characterize the diverse collection of mRNA binding proteins in the cell (Castello et al., 2012; Van Nostrand et al., 2017), we still have very little idea as to the composition of a single mRNP, and how this composition varies from mRNA to mRNA (or even varies between two mRNAs encoding the same gene). A more thorough investigation using both proteomic and single-molecule approaches to define an "mRNP-code" will provide insight into the communication between the ribosome, translation factors, and these mRNPs as they relate to translation.

Finally, although the phase of ribosome recycling (Figure 5) is reasonably well-determined biochemically and structurally, potential connections between recycling and re-initiation through ABCE1 remain particularly exciting. Although there is some evidence to suggest a role for ABCE1 in connecting these two translation events, how ABCE1 works to recruit initiation factors and promote pre-initiation complex assembly remains to be determined. Given all of these areas for further investigation, and the

constant discovery of new regulatory factors and sequence elements that modulate translation, this field will continue to be an exciting area for scientific exploration.

FIGURE LEGENDS

Figure 1. Overview of the eukaryotic translation cycle

The general process of translation begins with initiation where a complex coordination of many initiation factors (eIFs), initiator tRNA (Met-tRNA^{iMet}), the ribosomal subunits and mRNA to be translated come together at the AUG start codon of the open reading frame. Next, elongation allows for synthesis of the peptide chain through the efforts of elongation factors (eEFs) and aminoacyl-tRNAs (aa-tRNAs) until the ribosome reaches a termination or stop codon. In this termination phase the peptide is released with the help of specialized release factors (eRFs). Finally, the ribosome subunits must be recycled for a subsequent round of translation by the recycling factor ABCE1.

Figure 2. Translation elongation and various ribosome stalling events

(A) Overview of translation elongation. Aminoacyl-tRNAs are delivered to the ribosome in complex with eEF1A and GTP. Subsequent translocation by eEF2 causes tRNA repositioning from a hybrid state to classical state, creating an open A site for the next aa-tRNA. (B) Ribosome stalling due to slow peptidyl-transfer kinetics is rescued by eIF5A. (C) The use of a rare codon in the mRNA can cause ribosome stalling which can lead to mis-incorporation or frameshifting. (D) Certain tRNAs/codon pair orders can ribosome cause ribosome stalling. (E) mRNA secondary structure elements can cause ribosome stalling and programmed frameshifting at slippery sequences upstream of the structure.

Figure 3. Translation termination the role of mRNA context

(A) Overview of translation termination. When the ribosome encounters a termination codon, release factor eRF1 is delivered by eRF3 to coordinate peptidyl-hydrolysis at the ribosome active site. (B) Comparison of peptidyl-transfer chemistry by aminoacyl-tRNAs and peptidyl-release by an eRF1-coordinated water molecule in the ribosome active site. (C) RNA-binding proteins (RBPs) such as PABP can affect translation termination and lead to efficient peptide release, or if they are not present can cause inefficient release which in-turn leads to near-cognate incorporation or frameshifting.

Figure 4. Connections between NMD and contextual termination

A potential common mechanism between context-based termination and NMD involves the location of poly(A)-binding protein (PABP) or other mRNA-binding proteins in relation to the termination codon. In certain contexts, the termination codon would lead to proper termination and peptidyl-release. In other contexts, such as increased distance to PABP or close-proximity to the EJC, termination is impaired and the mRNA subjected to NMD.

Figure 5. Ribosome recycling and rescue

(A) Overview of ribosome recycling by ABCE1. ABCE1 binds to 80S ribosomes loaded with eRF1 and uses the power of ATP-binding and hydrolysis to dissociate ribosomal subunits. ABCE1 remains bound to the 40S subunit where it functions to stimulate subsequent translation initiation steps. (B) Ribosome rescue by ABCE1:Pelota releases stalled ribosomes due to endonucleolytic cleavage or ribosomes found in 3'UTRs. (C) Inactive, Stm1-bound 80S ribosomes can be rescued by ABCE1:Pelota to re-enter the cytoplasmic pool of translating ribosomes.

Figure 6. Recycling defects and implications in human disease

Dynamic regulation of ribosome rescue factors ABCE1 and Pelota is critical for ribosome homeostasis and translational output in blood cell lineages. During the initial phase of ABCE1 loss as the cells differentiate toward erythrocytes, the cells quickly upregulate the ribosome rescue factor Pelota which allows for increased recycling of the 3'UTR ribosomes, and translational reprogramming to synthesize hemoglobin. As the cells continue differentiation, Pelota:Hbs1L levels decrease over time from standard protein degradation pathways, 3'UTR ribosomes are again accumulated, and this 3'UTR accumulation eventually leads to a global translation defect and a trend toward decreased hemoglobin output.

REFERENCES

Alkalaeva, E.Z., Pisarev, A.V., Frolova, L.Y., Kisselev, L.L., and Pestova, T.V. (2006). In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. Cell *125*, 1125-1136.

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112-118.

Andersen, C.B., Becker, T., Blau, M., Anand, M., Halic, M., Balar, B., Mielke, T., Boesen, T., Pedersen, J.S., Spahn, C.M.*, et al.* (2006). Structure of eEF3 and the mechanism of transfer RNA release from the E-site. Nature *443*, 663-668.

Andersen, D.S., and Leevers, S.J. (2007). The essential Drosophila ATP-binding cassette domain protein, pixie, binds the 40 S ribosome in an ATP-dependent manner and is required for translation initiation. The Journal of biological chemistry *282*, 14752-14760.

Arribere, J.A., Cenik, E.S., Jain, N., Hess, G.T., Lee, C.H., Bassik, M.C., and Fire, A.Z. (2016). Translation readthrough mitigation. Nature *534*, 719-723.

Barthelme, D., Dinkelaker, S., Albers, S.V., Londei, P., Ermler, U., and Tampe, R. (2011). Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. Proceedings of the National Academy of Sciences of the United States of America *108*, 3228-3233.

Belew, A.T., Meskauskas, A., Musalgaonkar, S., Advani, V.M., Sulima, S.O., Kasprzak, W.K., Shapiro, B.A., and Dinman, J.D. (2014). Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. Nature *512*, 265-269.

Benne, R., and Hershey, J.W. (1978). The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. The Journal of biological chemistry *253*, 3078-3087.

Biswas, P., Jiang, X., Pacchia, A.L., Dougherty, J.P., and Peltz, S.W. (2004). The human immunodeficiency virus type 1 ribosomal frameshifting site is an invariant sequence determinant and an important target for antiviral therapy. Journal of virology *78*, 2082-2087.

Bonetti, B., Fu, L., Moon, J., and Bedwell, D.M. (1995). The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in Saccharomyces cerevisiae. Journal of molecular biology *251*, 334-345.

Brown, A., Shao, S., Murray, J., Hegde, R.S., and Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. Nature *524*, 493-496.

Brown, C.M., Stockwell, P.A., Trotman, C.N., and Tate, W.P. (1990). Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. Nucleic acids research *18*, 6339-6345.

Caliskan, N., Wohlgemuth, I., Korniy, N., Pearson, M., Peske, F., and Rodnina, M.V. (2017). Conditional Switch between Frameshifting Regimes upon Translation of dnaX mRNA. Molecular cell *66*, 558-567 e554.

Carr-Schmid, A., Pfund, C., Craig, E.A., and Kinzy, T.G. (2002). Novel G-protein complex whose requirement is linked to the translational status of the cell. Molecular and cellular biology *22*, 2564-2574.

Carvalho, M.D., Carvalho, J.F., and Merrick, W.C. (1984). Biological characterization of various forms of elongation factor 1 from rabbit reticulocytes. Archives of biochemistry and biophysics *234*, 603-611.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M*., et al.* (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell *149*, 1393-1406.

Chen, L., Muhlrad, D., Hauryliuk, V., Cheng, Z., Lim, M.K., Shyp, V., Parker, R., and Song, H. (2010). Structure of the Dom34-Hbs1 complex and implications for no-go decay. Nature structural & molecular biology *17*, 1233-1240.

Chen, Z.Q., Dong, J., Ishimura, A., Daar, I., Hinnebusch, A.G., and Dean, M. (2006). The essential vertebrate ABCE1 protein interacts with eukaryotic initiation factors. The Journal of biological chemistry *281*, 7452-7457.

Dana, A., and Tuller, T. (2014). The effect of tRNA levels on decoding times of mRNA codons. Nucleic acids research *42*, 9171-9181.

des Georges, A., Hashem, Y., Unbehaun, A., Grassucci, R.A., Taylor, D., Hellen, C.U., Pestova, T.V., and Frank, J. (2014). Structure of the mammalian ribosomal pre-termination complex associated with eRF1.eRF3.GDPNP. Nucleic acids research *42*, 3409-3418.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harbor perspectives in biology *4*, a013706.

Dinman, J.D. (2012). Mechanisms and implications of programmed translational frameshifting. Wiley interdisciplinary reviews RNA *3*, 661-673.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85-88.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561-564.

Dong, J., Lai, R., Nielsen, K., Fekete, C.A., Qiu, H., and Hinnebusch, A.G. (2004). The essential ATP-binding cassette protein RLI1 functions in translation by promoting preinitiation complex assembly. The Journal of biological chemistry *279*, 42157-42168.

dos Reis, M., Savva, R., and Wernisch, L. (2004). Solving the riddle of codon usage preferences: a test for translational selection. Nucleic acids research *32*, 5036-5044.

Draptchinskaia, N., Gustavsson, P., Andersson, B., Pettersson, M., Willig, T.N., Dianzani, I., Ball, S., Tchernia, G., Klar, J., Matsson, H*., et al.* (1999). The gene encoding ribosomal protein S19 is mutated in Diamond-Blackfan anaemia. Nature genetics *21*, 169-175.

Elf, J., Nilsson, D., Tenson, T., and Ehrenberg, M. (2003). Selective charging of tRNA isoacceptors explains patterns of codon usage. Science *300*, 1718-1722.

Eyler, D.E., Wehner, K.A., and Green, R. (2013). Eukaryotic release factor 3 is required for multiple turnovers of peptide release catalysis by eukaryotic release factor 1. The Journal of biological chemistry *288*, 29530-29538.

Ferguson, A., Wang, L., Altman, R.B., Terry, D.S., Juette, M.F., Burnett, B.J., Alejo, J.L., Dass, R.A., Parks, M.M., Vincent, C.T*., et al.* (2015). Functional Dynamics within the Human Ribosome Regulate the Rate of Active Protein Synthesis. Molecular cell *60*, 475-486.

Fischer, N., Neumann, P., Konevega, A.L., Bock, L.V., Ficner, R., Rodnina, M.V., and Stark, H. (2015). Structure of the E. coli ribosome-EF-Tu complex at <3 A resolution by Cs-corrected cryo-EM. Nature *520*, 567-570.

Floquet, C., Hatin, I., Rousset, J.P., and Bidou, L. (2012). Statistical analysis of readthrough levels for nonsense mutations in mammalian cells reveals a major determinant of response to gentamicin. PLoS genetics *8*, e1002608.

Frolova, L., Le Goff, X., Rasmussen, H.H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A.L., Celis, J.E., Philippe, M*., et al.* (1994). A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. Nature *372*, 701-703.

Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. Rna *2*, 334-341.

Frolova, L.Y., Tsivkovskii, R.Y., Sivolobova, G.F., Oparina, N.Y., Serpinsky, O.I., Blinov, V.M., Tatkov, S.I., and Kisselev, L.L. (1999). Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. Rna *5*, 1014-1020.

Gallant, J.A., and Lindsley, D. (1998). Ribosomes can slide over and beyond "hungry" codons, resuming protein chain elongation many nucleotides downstream. Proceedings of the National Academy of Sciences of the United States of America *95*, 13771-13776.

Gamble, C.E., Brule, C.E., Dean, K.M., Fields, S., and Grayhack, E.J. (2016). Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. Cell *166*, 679-690.

Gardin, J., Yeasmin, R., Yurovsky, A., Cai, Y., Skiena, S., and Futcher, B. (2014). Measurement of average decoding rates of the 61 sense codons in vivo. eLife *3*.

Glick, B.R., Chladek, S., and Ganoza, M.C. (1979). Peptide bond formation stimulated by protein synthesis factor EF-P depends on the aminoacyl moiety of the acceptor. European journal of biochemistry *97*, 23-28.

Glick, B.R., and Ganoza, M.C. (1975). Identification of a soluble protein that stimulates peptide bond synthesis. Proceedings of the National Academy of Sciences of the United States of America *72*, 4257-4260.

Gnirke, A., Geigenmuller, U., Rheinberger, H.J., and Nierhaus, L.H. (1989). The allosteric three-site model for the ribosomal elongation cycle. Analysis with a heteropolymeric mRNA. The Journal of biological chemistry *264*, 7291-7301.

Graille, M., Chaillet, M., and van Tilbeurgh, H. (2008). Structure of yeast Dom34: a protein related to translation termination factor Erf1 and involved in No-Go decay. The Journal of biological chemistry *283*, 7145-7154.

Graille, M., and Seraphin, B. (2012). Surveillance pathways rescuing eukaryotic ribosomes lost in translation. Nature reviews Molecular cell biology *13*, 727-735.

Gregio, A.P., Cano, V.P., Avaca, J.S., Valentini, S.R., and Zanelli, C.F. (2009). eIF5A has a function in the elongation step of translation in yeast. Biochemical and biophysical research communications *380*, 785-790.

Gurvich, O.L., Baranov, P.V., Gesteland, R.F., and Atkins, J.F. (2005). Expression levels influence ribosomal frameshifting at the tandem rare arginine codons AGG_AGG and AGA_AGA in Escherichia coli. Journal of bacteriology *187*, 4023-4032.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Molecular cell *51*, 35-45.

Guydosh, N.R., and Green, R. (2014). Dom34 rescues ribosomes in 3' untranslated regions. Cell *156*, 950-962.

Guydosh, N.R., and Green, R. (2017). Translation of poly(A) tails leads to precise mRNA cleavage. Rna *23*, 749-761.

Hanson, G., and Coller, J. (2017). Codon optimality, bias and usage in translation and mRNA decay. Nat Rev Mol Cell Biol, http://dx.doi.org/10.1038/nrm.2017.1091

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., and Jacobson, A. (1993). Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. Proceedings of the National Academy of Sciences of the United States of America *90*, 7034-7038.

Heaphy, S.M., Mariotti, M., Gladyshev, V.N., Atkins, J.F., and Baranov, P.V. (2016). Novel Ciliate Genetic Code Variants Including the Reassignment of All Three Stop Codons to Sense Codons in Condylostoma magnum. Molecular biology and evolution *33*, 2885-2889.

Henderson, A., and Hershey, J.W. (2011). Eukaryotic translation initiation factor (eIF) 5A stimulates protein synthesis in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America *108*, 6415-6419.

Heuer, A., Gerovac, M., Schmidt, C., Trowitzsch, S., Preis, A., Kotter, P., Berninghausen, O., Becker, T., Beckmann, R., and Tampe, R. (2017). Structure of the 40S-ABCE1 post-splitting complex in ribosome recycling and translation initiation. Nature structural & molecular biology *24*, 453-460.

Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. Annual review of biochemistry *83*, 779-812.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harbor perspectives in biology *4*.

Holbrook, J.A., Neu-Yilik, G., Hentze, M.W., and Kulozik, A.E. (2004). Nonsense-mediated decay approaches the clinic. Nature genetics *36*, 801-808.

Hoshino, S., Imai, M., Kobayashi, T., Uchida, N., and Katada, T. (1999). The eukaryotic polypeptide chain releasing factor (eRF3/GSPT) carrying the translation termination signal to the 3'-Poly(A) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein. The Journal of biological chemistry *274*, 16677-16680.

Ikemura, T., and Ozeki, H. (1983). Codon usage and transfer RNA contents: organism-specific codon-choice patterns in reference to the isoacceptor contents. Cold Spring Harbor symposia on quantitative biology *47 Pt 2*, 1087-1097.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Ivanov, A., Mikhailova, T., Eliseev, B., Yeramala, L., Sokolova, E., Susorov, D., Shuvalov, A., Schaffitzel, C., and Alkalaeva, E. (2016). PABP enhances release factor recruitment and stop codon recognition during translation termination. Nucleic acids research *44*, 7766-7776.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. The EMBO journal *27*, 736-747.

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., and Varmus, H.E. (1988). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature *331*, 280-283.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nature reviews Molecular cell biology *11*, 113-127.

Kane, J.F. (1995). Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Current opinion in biotechnology *6*, 494-500.

Karcher, A., Schele, A., and Hopfner, K.P. (2008). X-ray structure of the complete ABC enzyme ABCE1 from Pyrococcus abyssi. The Journal of biological chemistry *283*, 7962-7971.

Kemper, W.M., Berry, K.W., and Merrick, W.C. (1976). Purification and properties of rabbit reticulocyte protein synthesis initiation factors M2Balpha and M2Bbeta. The Journal of biological chemistry *251*, 5551-5557.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nature reviews Molecular cell biology *13*, 700-712.

Kong, C., Ito, K., Walsh, M.A., Wada, M., Liu, Y., Kumar, S., Barford, D., Nakamura, Y., and Song, H. (2004). Crystal structure and functional analysis of the eukaryotic class II release factor eRF3 from S. pombe. Molecular cell *14*, 233-245.

Kryuchkova, P., Grishin, A., Eliseev, B., Karyagina, A., Frolova, L., and Alkalaeva, E. (2013). Two-step model of stop codon recognition by eukaryotic release factor eRF1. Nucleic acids research *41*, 4573-4586.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nature methods *11*, 319-324.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. The EMBO journal *20*, 4987-4997.

Lee, H.H., Kim, Y.S., Kim, K.H., Heo, I., Kim, S.K., Kim, O., Kim, H.K., Yoon, J.Y., Kim, H.S., Kim, D.J*., et al.* (2007). Structural and functional insights into Dom34, a key component of no-go mRNA decay. Molecular cell *27*, 938-950.

Lobanov, A.V., Heaphy, S.M., Turanov, A.A., Gerashchenko, M.V., Pucciarelli, S., Devaraj, R.R., Xie, F., Petyuk, V.A., Smith, R.D., Klobutcher, L.A*., et al.* (2017).

Position-dependent termination and widespread obligatory frameshifting in Euplotes translation. Nature structural & molecular biology *24*, 61-68.

Mancera-Martinez, E., Brito Querido, J., Valasek, L.S., Simonetti, A., and Hashem, Y. (2017). ABCE1: A special factor that orchestrates translation at the crossroad between recycling and initiation. RNA biology, 1-7.

Marquardt, S., Hazelbaker, D.Z., and Buratowski, S. (2011). Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. Transcription *2*, 145-154.

Matheisl, S., Berninghausen, O., Becker, T., and Beckmann, R. (2015). Structure of a human translation termination complex. Nucleic acids research *43*, 8615-8626.

Melero, R., Buchwald, G., Castano, R., Raabe, M., Gil, D., Lazaro, M., Urlaub, H., Conti, E., and Llorca, O. (2012). The cryo-EM structure of the UPF-EJC complex shows UPF1 poised toward the RNA 3' end. Nature structural & molecular biology *19*, 498-505, S491-492.

Melnikov, S., Mailliot, J., Shin, B.S., Rigger, L., Yusupova, G., Micura, R., Dever, T.E., and Yusupov, M. (2016). Crystal Structure of Hypusine-Containing Translation Factor eIF5A Bound to a Rotated Eukaryotic Ribosome. Journal of molecular biology *428*, 3570-3576.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nature genetics *36*, 1073-1078.

Mills, E.W., Wangen, J., Green, R., and Ingolia, N.T. (2016). Dynamic Regulation of a Ribosome Rescue Pathway in Erythroid Cells and Platelets. Cell reports *17*, 1-10.

Mitrovich, Q.M., and Anderson, P. (2000). Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans. Genes & development *14*, 2173-2184.

Moazed, D., and Noller, H.F. (1989). Intermediate states in the movement of transfer RNA in the ribosome. Nature *342*, 142-148.

Namy, O., Moran, S.J., Stuart, D.I., Gilbert, R.J., and Brierley, I. (2006). A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. Nature *441*, 244-247.

Narla, A., and Ebert, B.L. (2010). Ribosomopathies: human disorders of ribosome dysfunction. Blood *115*, 3196-3205.

Nurenberg, E., and Tampe, R. (2013). Tying up loose ends: ribosome recycling in eukaryotes and archaea. Trends in biochemical sciences *38*, 64-74.

Pape, T., Wintermeyer, W., and Rodnina, M.V. (1998). Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome. The EMBO journal *17*, 7490-7497.

Park, M.H., Cooper, H.L., and Folk, J.E. (1981). Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. Proceedings of the National Academy of Sciences of the United States of America *78*, 2869-2873.

Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M., and Forster, A.C. (2009). Slow peptide bond formation by proline and other N-alkylamino acids in translation. Proceedings of the National Academy of Sciences of the United States of America *106*, 50-54.

Pechmann, S., and Frydman, J. (2013). Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. Nature structural & molecular biology *20*, 237-243.

Pelechano, V., and Alepuz, P. (2017). eIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences. Nucleic acids research *45*, 7326-7338.

Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Molecular cell *37*, 196-210.

Pisareva, V.P., Skabkin, M.A., Hellen, C.U., Pestova, T.V., and Pisarev, A.V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. The EMBO journal *30*, 1804-1817.

Plotkin, J.B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. Nature reviews Genetics *12*, 32-42.

Popp, M.W., and Maquat, L.E. (2013). Organizing principles of mammalian nonsense-mediated mRNA decay. Annual review of genetics *47*, 139-165.

Preis, A., Heuer, A., Barrio-Garcia, C., Hauser, A., Eyler, D.E., Berninghausen, O., Green, R., Becker, T., and Beckmann, R. (2014). Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. Cell reports *8*, 59-65.

Quax, T.E., Claassens, N.J., Soll, D., and van der Oost, J. (2015). Codon Bias as a Means to Fine-Tune Gene Expression. Molecular cell *59*, 149-161.

Rodnina, M.V., Savelsbergh, A., Katunin, V.I., and Wintermeyer, W. (1997). Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. Nature *385*, 37-41.

Rossi, D., Barbosa, N.M., Galvao, F.C., Boldrin, P.E., Hershey, J.W., Zanelli, C.F., Fraser, C.S., and Valentini, S.R. (2016). Evidence for a Negative Cooperativity between eIF5A and eEF2 on Binding to the Ribosome. PloS one *11*, e0154205.

Roy, B., Leszyk, J.D., Mangus, D.A., and Jacobson, A. (2015). Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3. Proceedings of the National Academy of Sciences of the United States of America *112*, 3038-3043.

Saini, P., Eyler, D.E., Green, R., and Dever, T.E. (2009). Hypusine-containing protein eIF5A promotes translation elongation. Nature *459*, 118-121.

Salas-Marco, J., and Bedwell, D.M. (2004). GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. Molecular and cellular biology *24*, 7769-7778.

Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2016). Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. Nucleic acids research *44*, 1944-1951.

Schreier, M.H., Erni, B., and Staehelin, T. (1977). Initiation of mammalian protein synthesis. I. Purification and characterization of seven initiation factors. Journal of molecular biology *116*, 727-753.

Schueren, F., and Thoms, S. (2016). Functional Translational Readthrough: A Systems Biology Perspective. PLoS genetics *12*, e1006196.

Schuller, A.P., Wu, C.C., Dever, T.E., Buskirk, A.R., and Green, R. (2017). eIF5A Functions Globally in Translation Elongation and Termination. Molecular cell *66*, 194-205 e195.

Sengupta, J., Nilsson, J., Gursky, R., Kjeldgaard, M., Nissen, P., and Frank, J. (2008). Visualization of the eEF2-80S ribosome transition-state complex by cryo-electron microscopy. Journal of molecular biology *382*, 179-187.

Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., and Hegde, R.S. (2016). Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. Cell *167*, 1229-1240 e1215.

Sharp, P.M., and Li, W.H. (1987). The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. Nucleic acids research *15*, 1281-1295.

Shemin, D., and Rittenberg, D. (1946). The life span of the human red blood cell. The Journal of biological chemistry *166*, 627-636.

Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. Science *330*, 369-372.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proceedings of the National Academy of Sciences of the United States of America *108*, E1392-1398.

Shoemaker, C.J., and Green, R. (2012). Translation drives mRNA quality control. Nature structural & molecular biology *19*, 594-601.

Simonetti, A., Brito Querido, J., Myasnikov, A.G., Mancera-Martinez, E., Renaud, A., Kuhn, L., and Hashem, Y. (2016). eIF3 Peripheral Subunits Rearrangement after mRNA Binding and Start-Codon Recognition. Molecular cell *63*, 206-217.

Smith, J.E., Alvarez-Dominguez, J.R., Kline, N., Huynh, N.J., Geisler, S., Hu, W., Coller, J., and Baker, K.E. (2014). Translation of small open reading frames within unannotated RNA transcripts in Saccharomyces cerevisiae. Cell reports *7*, 1858-1866.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Song, H., Mugnier, P., Das, A.K., Webb, H.M., Evans, D.R., Tuite, M.F., Hemmings, B.A., and Barford, D. (2000). The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. Cell *100*, 311-321.

Swart, E.C., Serra, V., Petroni, G., and Nowacki, M. (2016). Genetic Codes with No Dedicated Stop Codon: Context-Dependent Translation Termination. Cell *166*, 691-702.

Taylor, D.J., Nilsson, J., Merrill, A.R., Andersen, G.R., Nissen, P., and Frank, J. (2007). Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation. The EMBO journal *26*, 2421-2431.

Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R.N., and Chrzanowska-Lightowlers, Z.M. (2010). Hungry codons promote frameshifting in human mitochondrial ribosomes. Science *327*, 301.

Triana-Alonso, F.J., Chakraburtty, K., and Nierhaus, K.H. (1995). The elongation factor 3 unique in higher fungi and essential for protein biosynthesis is an E site factor. The Journal of biological chemistry *270*, 20473-20478.

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. Molecular cell *46*, 518-529.

Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., and Jung, K. (2013). Translation Elongation Factor EF-P Alleviates Ribosome Stalling at Polyproline Stretches. Science *339*, 82-85.

van den Elzen, A.M., Schuller, A., Green, R., and Seraphin, B. (2014). Dom34-Hbs1 mediated dissociation of inactive 80S ribosomes promotes restart of translation after stress. The EMBO journal *33*, 265-276.

Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Blue, S.M., Dominguez, D., Cody, N.A.L., Olson, S., Sundararaman, B.*, et al.* (2017). A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. bioRxiv *179648*, http://dx.doi.org/10.1101/179648.

Voorhees, R.M., and Ramakrishnan, V. (2013). Structural basis of the translational elongation cycle. Annual review of biochemistry *82*, 203-236.

Welch, E.M., and Jacobson, A. (1999). An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA. The EMBO journal *18*, 6134-6145.

Wilson, D.N., Arenz, S., and Beckmann, R. (2016). Translation regulation via nascent polypeptide-mediated ribosome stalling. Current opinion in structural biology *37*, 123-133.

Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M.V. (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. The Journal of biological chemistry *283*, 32229-32235.

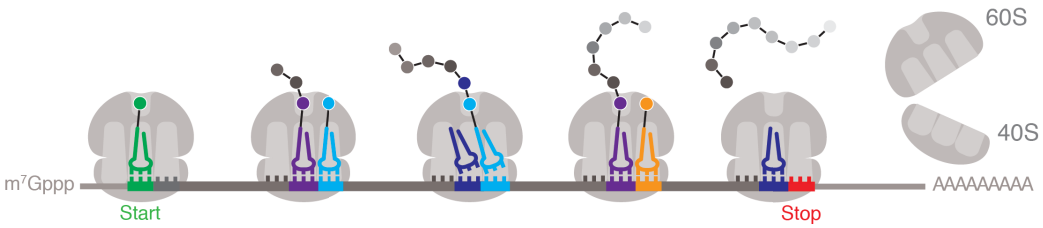
Yan, S., Wen, J.D., Bustamante, C., and Tinoco, I., Jr. (2015). Ribosome excursions during mRNA translocation mediate broad branching of frameshift pathways. Cell *160*, 870-881.

Young, D.J., Guydosh, N.R., Zhang, F., Hinnebusch, A.G., and Green, R. (2015). Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation Reinitiation in 3'UTRs In Vivo. Cell *162*, 872-884.

Youngman, E.M., Brunelle, J.L., Kochaniak, A.B., and Green, R. (2004). The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. Cell *117*, 589-599.

Yu, C.H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M.S., and Liu, Y. (2015). Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. Molecular cell *59*, 744-754.

Zahonova, K., Kostygov, A.Y., Sevcikova, T., Yurchenko, V., and Elias, M. (2016). An Unprecedented Non-canonical Nuclear Genetic Code with All Three Termination Codons Reassigned as Sense Codons. Current biology : CB *26*, 2364-2369.



Phase:	Initiation	Elongation	Termination	Recycling
Factors:	eIFs Met-tRNA^{iMet}	eEFs aa-tRNAs	eRFs	ABCE1

Figure 1

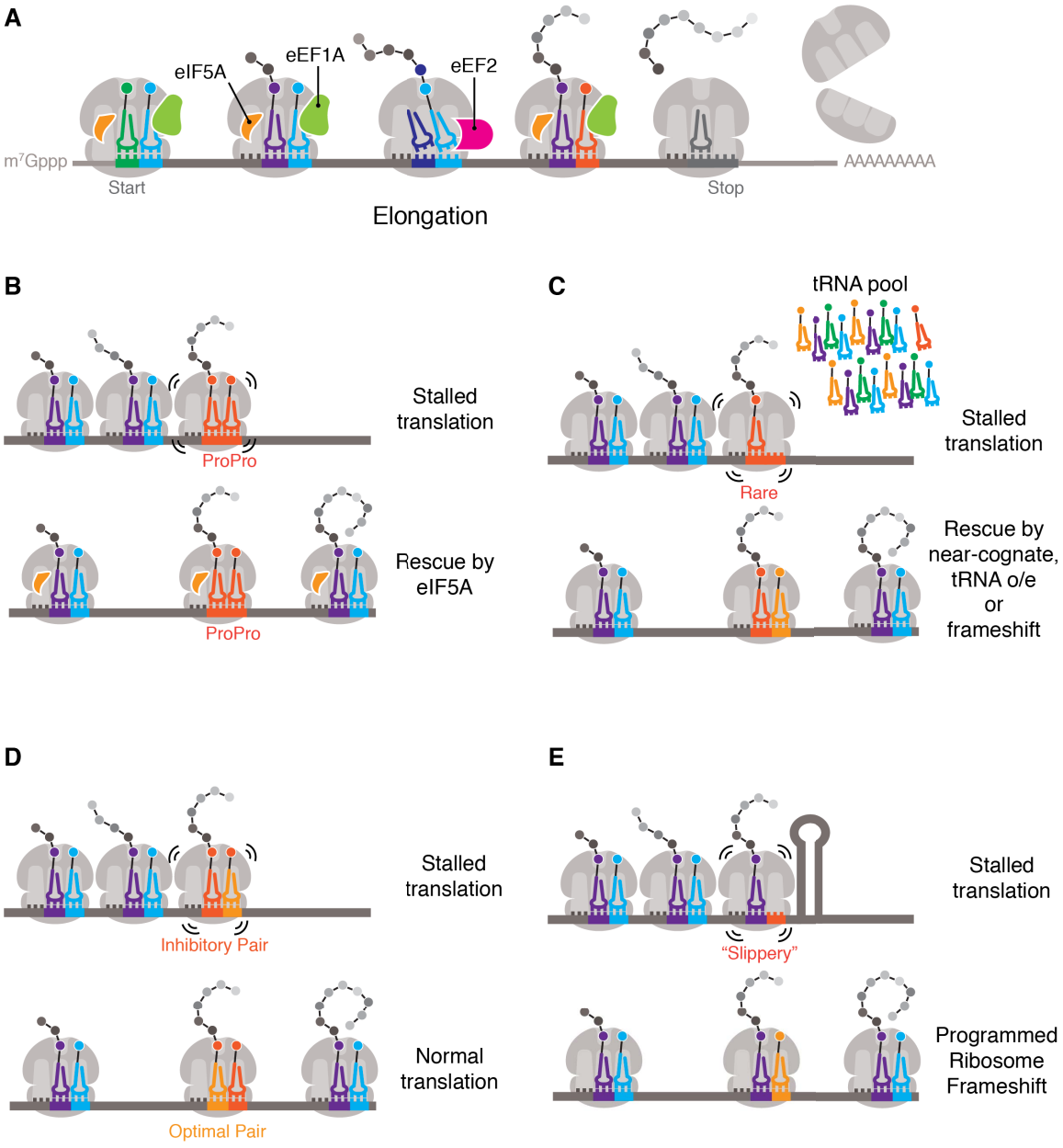

A
eIF5A
eEF1A
eEF2
m7Gppp
Start
Stop
AAAAAAAAA
Elongation
B
ProPro
Stalled translation
ProPro
Rescue by eIF5A
C
tRNA pool
Rare
Stalled translation
Rescue by near-cognate, tRNA o/e or frameshift
D
Inhibitory Pair
Stalled translation
Optimal Pair
Normal translation
E
"Slippery"
Stalled translation
Programmed Ribosome Frameshift

Figure 2

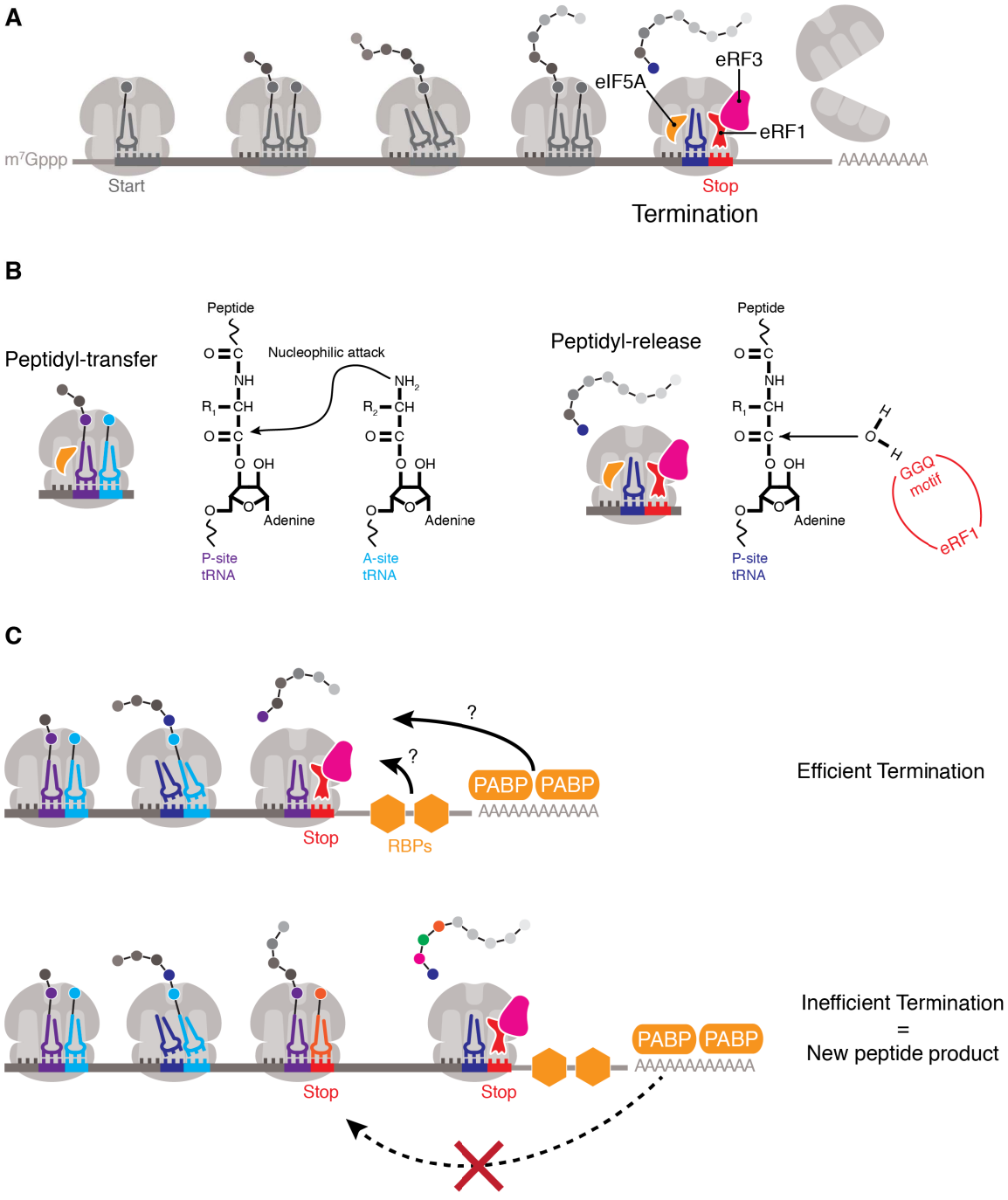

A
eIF5A
eRF3
eRF1
m7Gppp
AAAAAAAAA
Start
Stop
Termination
B
Peptidyl-transfer
Peptide
Nucleophilic attack
Adenine
P-site
tRNA
A-site
tRNA
Peptidyl-release
GGQ
motif
eRF1
C
?
?
PABP
PABP
AAAAAAAAAAA
Stop
RBPs
Efficient Termination
Inefficient Termination
=
New peptide product

Figure 3

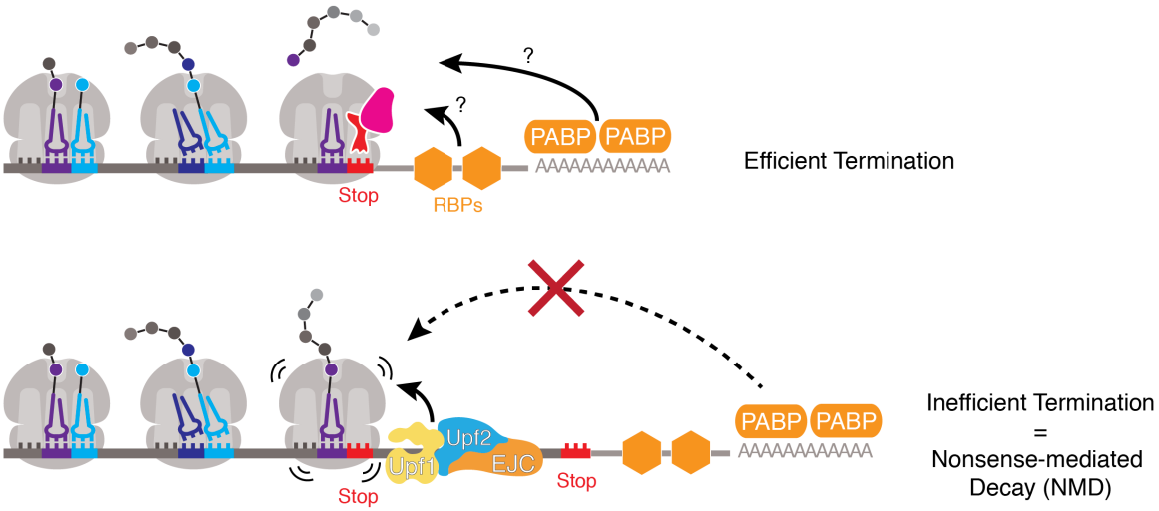

?
?
PABP
PABP
AAAAAAAAAAA
Stop
RBPs
Efficient Termination
PABP
PABP
AAAAAAAAAAA
Upf2
Upf1
EJC
Stop
Stop
Inefficient Termination
=
Nonsense-mediated
Decay (NMD)

Figure 4

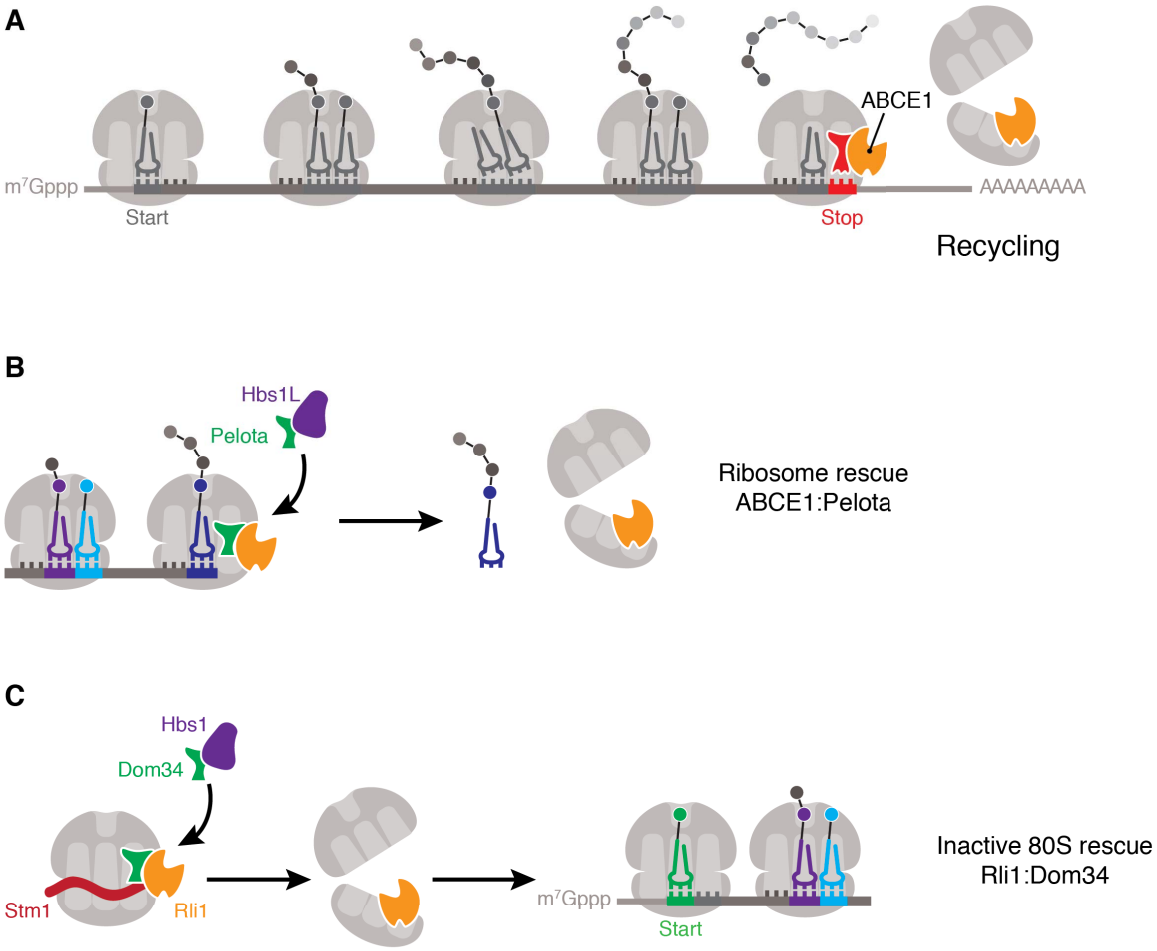

A
m7Gppp
Start
ABCE1
Stop
AAAAAAAAA
Recycling
B
Hbs1L
Pelota
Ribosome rescue
ABCE1:Pelota
C
Hbs1
Dom34
Stm1
Rli1
m7Gppp
Start
Inactive 80S rescue
Rli1:Dom34

Figure 5

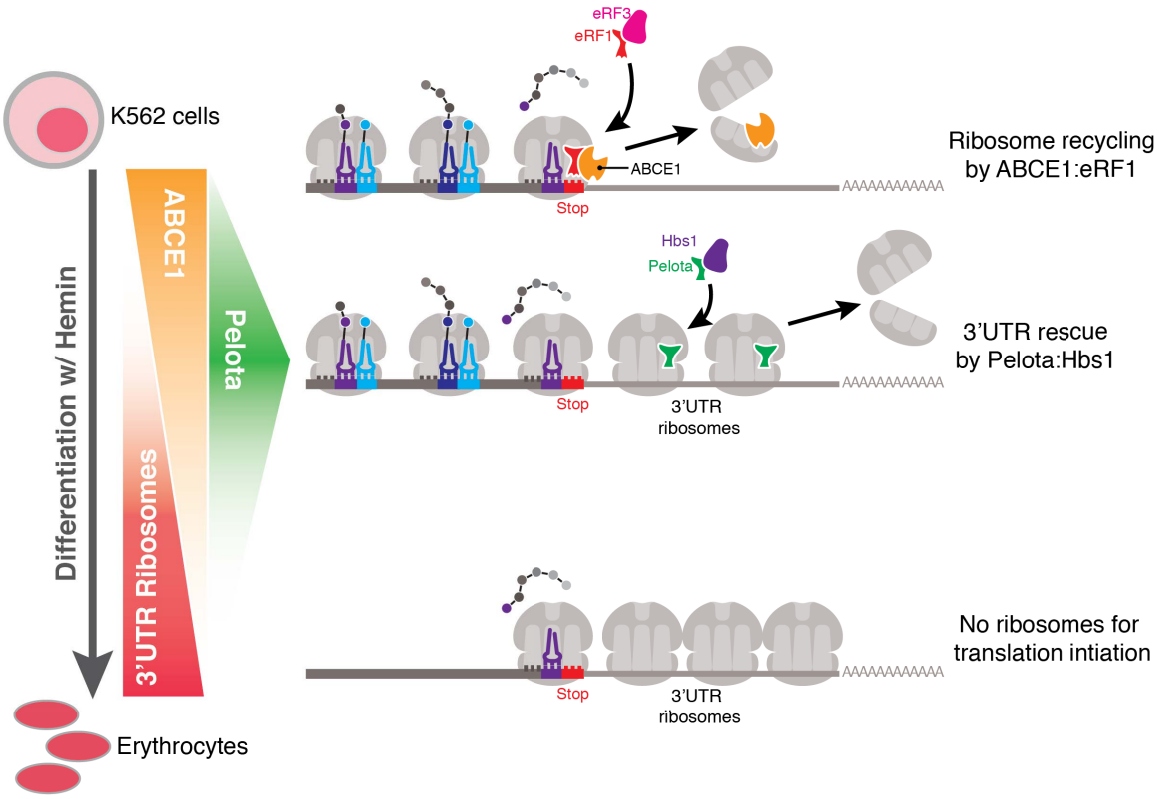

K562 cells
Differentiation w/ Hemin
ABCE1
Pelota
3'UTR Ribosomes
Erythrocytes
eRF3
eRF1
ABCE1
Stop
AAAAAAAAAAA
Ribosome recycling
by ABCE1:eRF1
Hbs1
Pelota
Stop
3'UTR
ribosomes
AAAAAAAAAAA
3'UTR rescue
by Pelota:Hbs1
Stop
3'UTR
ribosomes
AAAAAAAAAAA
No ribosomes for
translation intiation

Figure 6

Chapter II: eIF5A stimulates global translation elongation and termination

ABSTRACT

The eukaryotic translation factor eIF5A, originally identified as an initiation factor, was later shown to promote translation elongation of iterated proline sequences. Using a combination of ribosome profiling and *in vitro* biochemistry, we report a much broader role for eIF5A in elongation and uncover a critical function for eIF5A in termination. Ribosome profiling of an eIF5A-depleted strain reveals a global elongation defect, with abundant ribosomes stalling at many sequences, not limited to proline stretches. Our data also show ribosome accumulation at stop codons and in the 3'-UTR, suggesting a global defect in termination in the absence of eIF5A. Using an *in vitro* reconstituted translation system, we find that eIF5A strongly promotes the translation of the stalling sequences identified by profiling and increases the rate of peptidyl-tRNA hydrolysis more than 17-fold. We conclude that eIF5A functions broadly in elongation and termination, rationalizing its high cellular abundance and essential nature.

This chapter was previously published in

Molecular Cell 66, 194-205 (2017)

INTRODUCTION

In addition to the core set of factors required for protein synthesis, many auxiliary proteins stimulate specific processes in the translation cycle. One such protein, eIF5A, was originally identified nearly 40 years ago as a factor that stimulates formation of the first peptide bond between Met-tRNA and puromycin (Benne and Hershey, 1978; Kemper et al., 1976; Schreier et al., 1977). More recent work from several groups identified eIF5A as having a more general role in elongation by conditionally depleting eIF5A from yeast cells *in vivo* and performing polysome analyses and ribosome transit measurements (Gregio et al., 2009; Henderson and Hershey, 2011; Saini et al., 2009). eIF5A is an essential gene in eukaryotes and the nature of its critical cellular function has been a subject of ongoing exploration.

eIF5A is a small, highly-expressed protein containing only 157 amino acids and is post-translationally modified with hypusine at a conserved lysine residue (Dever et al., 2014; Park et al., 1981). This hypusine modification is critical for eIF5A function *in vivo* and in assays of Met-Puromycin formation (Park et al., 2011; Park, 1989; Park et al., 1991; Saini et al., 2009). Recent biochemical work with the bacterial homolog of eIF5A, called EFP, revealed that EFP functions to promote the translation of polyproline containing peptides that stall the ribosome (in addition to Met-Puromycin formation) (Doerfel et al., 2013; Glick et al., 1979; Glick and Ganoza, 1975; Ude et al., 2013). Proline residues were shown to be poor substrates for peptide bond formation, likely due to the unique geometry that they assume, thereby leading to slower ribosome elongation (Pavlov et al., 2009; Wohlgemuth et al., 2008). Subsequent work showed that eIF5A similarly stimulates translation of polyproline peptides (with as few as two prolines) in

eukaryotes and that this function is highly dependent on its hypusine modification (Gutierrez et al., 2013).

While EFP and eIF5A appear to have similar biochemical functions in translation elongation, the essential nature of eIF5A in eukaryotic cells remains poorly understood. Given the documented role in polyproline synthesis and that approximately 10% of yeast genes contain stretches of three or more consecutive proline residues (and 95 proteins in yeast contain four or more consecutive Pro residues as compared with only 9 in *E. coli* (Doerfel et al., 2013)), it was rationalized that the global defect in elongation observed upon eIF5A depletion (Saini et al., 2009) resulted from defects in translation mediated by polyproline motifs (Gutierrez et al., 2013). These bioinformatic data led to the suggestion that eIF5A is essential in eukaryotes because of the relative abundance of polyproline motifs in these organisms.

Here we explore the *in vivo* and *in vitro* function of eIF5A in both translation elongation and termination. Through a combination of ribosome profiling and biochemistry in a reconstituted translation system, we find that eIF5A has a broader role in elongation than previously understood, and we uncover a critical function for eIF5A in translation termination. Importantly, we show that eIF5A functions to stimulate translation elongation in many peptide contexts, certainly not limited to proline stretches, and to accelerate the rate of peptidyl-tRNA hydrolysis by eRF1 during termination. Our findings suggest that eIF5A is an obligate translation factor acting on many (if not all) translating ribosomes, thereby rationalizing its essential nature and high abundance in eukaryotic cells.

RESULTS

Conditional depletion of eIF5A by an auxin-inducible degron

Given that eIF5A is encoded by an essential gene, we had to develop a method for conditional depletion in order to observe its *in vivo* function. Initial efforts to perform ribosome profiling with a temperature-sensitive allele were largely unsuccessful due to poor reproducibility and global changes in gene expression introduced by the temperature shift. Not only does the shift to the non-permissive temperature alter the transcriptional landscape, we suspected that it might affect translation elongation in a way that could confound analyses of protein synthesis (Grousl et al., 2013).

We constructed a yeast strain in which both the transcription of eIF5A and degradation of the protein can be manipulated by inducers added to the growth media. In this strain, eIF5A is expressed from the *GAL1* promoter; transcription levels are high in the presence of galactose and low in the presence of glucose. To promote rapid turnover of the protein, we fused eIF5A with a mini auxin-inducible degron (mAID) tag; in the presence of auxin, this domain recruits the E3-ubiquitin ligase TIR1 which ubiquitinates the mAID-eIF5A fusion protein for proteasome degradation (Nishimura and Kanemaki, 2014).

By inhibiting transcription and promoting turnover of the protein, this strategy effectively depletes eIF5A and broadly recapitulates observations using other methods (Gutierrez et al., 2013; Saini et al., 2009). Although the eIF5A-degron fusion protein (eIF5Ad) is expressed in our system at somewhat lower levels than endogenous eIF5A, it supports cell viability (Figure S1A and S1B). Following a switch to glucose media to shut off transcription and the addition of auxin to deplete eIF5A, cell growth arrests after

~10 h (Figure S1C). The timing of this arrest is consistent with the expectation that eIF5A is essential for cell growth: after ~8 h, we were unable to detect the eIF5A protein on immunoblots (Figure S1A). For additional validation, we analyzed polysome profiles and observed that after 10 h of induction, eIF5Ad cells showed an increase in the polysome/monosome ratio compared to WT cells (Figure S1D), as previously reported (Saini et al., 2009).

eIF5A depletion causes redistribution of ribosomes toward the 5' end of genes

We generated libraries of ribosome footprints from the WT and eIF5Ad strains after 10 h of conditional growth. We did not pre-treat the culture with antibiotics to arrest translation (Ingolia et al., 2009) because adding cycloheximide to the media has been shown to cause sequence-specific ribosome pausing *in vivo* (Hussmann et al., 2015). Instead, lysates were flash frozen in liquid nitrogen, lysed, and thawed in lysis buffer containing cycloheximide to prevent lingering elongation during lysis (Guydosh and Green, 2014; Weinberg et al., 2016). We generated two biological replicates of the WT and eIF5Ad libraries and obtained ~15-25 million mapped reads for each. The number of ribosome footprints per gene is highly reproducible between biological replicates (Pearson's r= 0.96 for eIF5Ad cells, 0.99 for WT cells) (Figure S1E).

Our previous observation that depletion of eIF5A increases the fraction of ribosomes in polysomes led to the argument that eIF5A is a general translation factor (Saini et al., 2009). To get a genome-wide view of this phenomenon in our ribosome profiling data, we plotted the average ribosome occupancy of all genes aligned at start codons. In eIF5Ad cells, we observed an increase in ribosome occupancy in the first ~50

codons relative to the wild-type strain and a decrease in downstream occupancy (Figure 1A). These trends are readily understood by examining specific genes where we observe higher ribosome occupancy in the 5'-end relative to the 3' end (Figure 1B). This phenomenon is consistent with a general elongation defect: an enrichment of ribosomes at the 5'-end of genes was previously observed during amino acid starvation or treatment with antibiotics that inhibit elongation, both of which result in global ribosome pausing (Gerashchenko and Gladyshev, 2014; Ingolia et al., 2009; Subramaniam et al., 2014).

To quantify the differences in the position of ribosomes along transcripts, we computed a polarity score for every gene (see Methods). This metric assigns a value between -1 and +1 to each gene based on the distribution of ribosome footprints along it. We excluded ribosome densities at both ends of genes (15 nucleotides) from our analysis to avoid known artifacts introduced by start and stop codons (Young et al., 2015). Enrichment of ribosome occupancy at the 5'-end of a gene gives a polarity score between -1 and 0 while enrichment toward the 3'-end of a gene results in a score between 0 and +1. This metric yields information about the balance of ribosome occupancy across the gene (5' to 3') but not about the fine details of the distribution. For example, two very different distributions of ribosome density could both yield a polarity score of 0: enrichment at the center of a gene or an equal enrichment at both ends. Nevertheless, we find that this metric reveals broad trends in our data that provide insight into eIF5A function. Strikingly, the distribution of ribosome occupancy in eIF5Ad cells is shifted to the left relative to the distribution in WT cells (Figures 1C and S2A). This quantitative metric nicely summarizes the skewed ribosome distribution observed in specific genes and in plots of average ribosome occupancy (Figures 1A and 1B).

The accumulation of ribosomes at the 5'-ends of genes in eIF5Ad cells is consistent with pauses in elongation followed by queuing of upstream ribosomes (as depicted in the model in Figure 1A). Given eIF5A's ability to resolve pausing at polyproline stretches, we hypothesized that the loss of eIF5A could explain these pauses in elongation. To test this idea, we divided yeast genes into two (roughly equal) subsets, one containing Pro-Pro dipeptide motifs and the other lacking these motifs. We chose to separate our data based on possession of Pro-Pro motifs because it was the minimal motif on which eIF5A was known to function (Gutierrez et al., 2013). We were surprised to find that both subsets showed a significant shift to the left, indicative of marked enrichment in ribosome occupancy at the 5'-end (Figures 1D and S2B), independent of the presence or absence of Pro-Pro motifs. This striking genome-wide observation of ribosome pausing suggests that eIF5A may functions broadly in eukaryotic cells, and may not simply relieve pausing at polyproline motifs.

eIF5A alleviates ribosome pausing at poly-Pro motifs

As a first step in analyzing translation pauses, we started with the known role of eIF5A in resolving pausing at polyproline stretches. We computed the average ribosome occupancy at two or three consecutive Pro codons in well-expressed genes and found the peak was ~2-fold higher in eIF5Ad cells than WT cells (Figures 2A and 2B). We observed two major peaks in the diproline plot at the center of the average gene plot (Figure 2A); one peak corresponds to pauses with the Pro codons positioned in the P and A sites while the second corresponds to pauses with the Pro codons positioned in the E and P sites. These observations suggest that forming the first Pro-Pro peptide bond is

challenging without eIF5A and that forming downstream peptide bonds is challenging when Pro-Pro dipeptide is engaged in the exit channel. These findings are consistent with previous biochemical and profiling studies of EFP in the *E. coli* system (Doerfel et al., 2013; Ude et al., 2013; Woolstenhulme et al., 2015).

Pauses at polyproline stretches affect the ribosome distribution surrounding the pausing motifs as well. We observed increased ribosome occupancy ~30 nt upstream of the polyproline motifs (Figures 2A and 2B, arrows) due to trailing ribosomes forming a queue behind those paused at the polyproline motif. In addition, ribosome occupancy downstream of polyproline motifs is strongly depleted, presumably because downstream ribosomes continue elongating, leaving empty mRNA behind them (Figure 2B, inset). This depletion is more pronounced for triproline motifs than diproline motifs, consistent with prior observations that a diproline motif does not reduce protein levels upon loss of eIF5A activity; an effect on protein expression is only seen with three or four consecutive Pro codons (Gutierrez et al., 2013). Our observations of strong pauses, stacked ribosomes upstream of the paused ribosome, and depletion of downstream density after PPP motifs are very similar to the effects seen in our earlier work upon deletion of the EFP gene in bacteria (Woolstenhulme et al., 2015).

We note that pausing at polyproline motifs is evident in WT cells, suggesting that these motifs are problematic even in the presence of factors that evolved to mitigate their impact on elongation (Pavlov et al., 2009; Wohlgemuth et al., 2008). Proline residues have been associated with ribosome pausing in several eukaryotic-based ribosome profiling studies (Artieri and Fraser, 2014; Ingolia et al., 2011), suggesting that the

endogenous activity of eIF5A may not be sufficient to wholly relieve ribosome pausing at polyproline motifs in these cells.

Widespread ribosome pausing in eIF5Ad cells

In our analyses of ribosome occupancy on individual genes, we also noticed many strong pauses in eIF5Ad cells at positions that do not encode polyproline motifs (e.g. *THO1* and *PCL6* in Figure 2C). These additional pauses may explain the polarity effects of eIF5A depletion on genes lacking polyproline motifs. We explored this phenomenon by computing pause scores for all 8,000 tripeptide motifs. In eIF5Ad cells, 445 tripeptide motifs showed at least a two-fold increase in pause scores compared to WT cells while 29 motifs have a pause score 10-fold or greater than the average density of the gene (Figure 2D, S3A and Table S1).

The top 29 tripeptide motifs show a consensus peptide sequence with enrichment of Pro or Asp in the E and the P sites and Pro in the A site; Gly is modestly enriched in all three sites (Figure 2D). Despite the importance of Pro in the consensus sequence, surprisingly 18 of these 29 motifs do not contain a Pro-Pro sequence. To demonstrate the impact of eIF5A on these identified sequences, we show an average gene analysis of ~5,500 sites corresponding to the 18 non Pro-Pro motifs that reveals a level of pausing upon eIF5A depletion (Figure 2E) comparable to that observed at PPP (Figure 2B). Moreover, strong pauses are observed at the four tripeptide motifs completely lacking Pro (DVG, DDG, GGT and RDK, Figure S3B). Taken together, our findings demonstrate that eIF5A has a much broader role in alleviating ribosome pausing than previously thought.

eIF5A promotes translation of stalling sequences in vitro

Ribosome profiling is not a time-resolved experiment and so there is no way of knowing whether all positions are impacted by the depletion of eIF5A. Profiling only reveals which motifs are more strongly impacted than others. To address this, we performed experiments *in vitro* using a reconstituted translation system to ask whether eIF5A is directly involved in promoting slow steps in translation elongation. We first tested the ability of our purified eIF5A to accelerate the reaction between Met-tRNA and the antibiotic puromycin since eIF5A was originally identified using this assay (Benne and Hershey, 1978; Kemper et al., 1976; Schreier et al., 1977). Pelleted 80S initiation complexes programmed with a simple mRNA (MFN-Stop) with [^{35}S]-Met-tRNA^{iMet} loaded in the ribosomal P site were mixed with puromycin (Puro) in the presence and absence of eIF5A, and formation of the Met-Puro product was observed over time. As anticipated, we find that post-translationally hypusinated eIF5A greatly enhances the rate of Met-Puro formation while unmodified eIF5A has a more subtle effect (Figure S4C). This experiment documents for the first time a >100-fold rate enhancement for this reaction resulting from the addition of modified eIF5A.

With this confirmation of robust *in vitro* activity, we tested the role of eIF5A in the translation of peptides containing several stalling motifs that were identified in our profiling experiments. While our profiling analysis described above focused on identifying tripeptide stalling motifs, we also are able to identify pausing motifs at the dipeptide level in our profiling data (Table S2) and because these motifs are simpler to characterize biochemically, we chose to focus efforts on these for the *in vitro* assays. In these experiments, pelleted 80S initiation complexes programmed with a simple mRNA

(MXXK-Stop where X is any amino acid) containing [^{35}S]-Met-tRNA^{iMet} were incubated with elongation factors and the appropriate tRNAs (eEF1A:aa-tRNA, eEF2, eEF3) and peptide formation was measured over time (Figure 3A). Each peptide contains the problematic dipeptide being analyzed (e.g. Pro-Pro) followed by a C-terminal lysine that helps us to resolve the tetrapeptide product using an electrophoretic TLC system.

As we previously reported (Gutierrez et al., 2013), eIF5A is essential for robust synthesis of a Pro-Pro containing peptide (MPPK) while having a negligible effect on a Phe-Phe containing peptide (MFFK). For the MPPK reaction, we see that the stimulatory effect of eIF5A is highly dependent on the hypusine modification (Figure 3B): unmodified eIF5A increases the reaction endpoint (amount of peptide formed, Y_{max}) but hypusinated eIF5A increases both the endpoint and the rate of the reaction (k_{obs}, compared to unmodified eIF5A). The endpoint defects are due to substantial peptidyl-tRNA dropoff in these slow peptidyl-transfer reactions (Figure S4D); similar trends for short peptidyl-tRNAs have been reported by others (Doerfel et al., 2013; Katoh et al., 2016). These data argue that the presence of unmodified eIF5A stabilizes the complex against peptidyl-tRNA dropoff but that the hypusine modification is critical for maximal rate enhancement.

We next evaluated the role of eIF5A in promoting elongation through several dipeptide stalling motifs identified from our ribosome profiling data (Table S2), focusing on hyspuinated eIF5A as this provided maximal stimulation of the Pro-Pro-containing peptide. In what follows, we limit our analysis to the endpoint of the reaction, given the difficulty of separating the rate for the reaction of interest from the competing dropoff reaction in observed rate constants. Since several of the strongest dipeptide stalling

motifs contained aspartic acid, we measured the synthesis of the MDDK and MDPK peptides with and without eIF5A (Figure 3C); eIF5A strongly promoted synthesis of both peptides. We then analyzed the effects of eIF5A for several other peptides containing combinations of proline, aspartic acid, cysteine, and phenylalanine. We anticipated that several of these motifs would induce strong pauses (DD, DP and PD) while others would induce little or no stalling (CC and CF), based on their relative pause scores (Table S2). Somewhat surprisingly, we find that eIF5A has a stimulatory effect on elongation for all the peptides tested (Figure 3D). This result points to the fact that our profiling data only measure relative pausing levels between peptide motifs, not absolute levels as are measured in our *in vitro* experiments.

While these experiments gave us insight into the breadth of stalling motifs that are rescued by eIF5A, we wanted to further investigate the role of hypusine in resolving these translation pauses. We chose to probe the requirement of eIF5A hypusine modification using several tripeptide motifs: a polyphenylalanine control (MFFFK), one with polyproline (MPPPK), one with a single proline (MPDIK), and one with no prolines (MDDIK). We find that eIF5A does not stimulate polyphenylalanine synthesis (MFFFK) while it is critical for polyproline synthesis (MPPPK) (Figures 4A and B). As observed for diproline (Figure 3B), non-hypusinated eIF5A increases the MPPPK reaction endpoint while hypusinated eIF5A is critical for maximal endpoint and rate enhancement. Interestingly, we find that hypusine is less important for translation of other non-polyproline stalling motifs, such as PDI and DDI (Figure 4C and D). Taken together, our *in vitro* and ribosome profiling data suggest that eIF5A plays a substantial role in the rate

of translation elongation at most (if not all) sites in the transcriptome, although hypusine modification may be required only for resolving polyproline stalls.

eIF5A promotes translation termination

Recent biochemical and structural studies have shown that eIF5A binds to the E site of the ribosome where it stabilizes the peptidyl-tRNA for nucleophilic attack by the incoming amino acid on the A-site tRNA (Gutierrez et al., 2013; Melnikov et al., 2016; Schmidt et al., 2016). We wondered if eIF5A also promotes peptidyl-tRNA hydrolysis by eRF1, given that both of these reactions would benefit from a well-ordered peptidyl-tRNA and active site geometry. Turning again to our profiling data, we computed the average ribosome occupancy of all genes aligned at stop codons (Figure 5A) and observed several major differences between the WT and eIF5Ad strains. First, the stop codon peak in eIF5Ad strain is ~2-fold higher than that seen in the WT strain; indeed, it is accompanied by a secondary peak ~30 nt upstream, presumably due to ribosomes stacked behind the terminating ribosome. Second, ribosome occupancy in the 3'-UTR is modestly increased (Figure 5B). Both of these observations are consistent with defects in translation termination in the eIF5Ad strain.

To ask whether ribosome occupancy at stop codons and in the 3'-UTR reflects elongating or scanning (post-termination) ribosomes, we treated lysates with a high salt buffer that releases ribosomes lacking a nascent polypeptide chain (Blobel and Sabatini, 1971; Mills et al., 2016). As expected, elongating ribosomes in coding regions and the stacked ribosomes located ~ 30 nt upstream of stop codons (in eIF5Ad cells) are not sensitive to the high salt wash. In contrast, the high salt wash released ~90% of

ribosomes at stop codons in WT cells (Figure 5C, top panel), suggesting that most ribosomes at stop codons have released their polypeptide chains and are waiting for the subunits to be split and recycled. On the other hand, only ~25% of ribosomes at stop codons were washed away in eIF5Ad cells (Figure 5C, bottom panel), suggesting that a majority of these ribosomes have yet to properly terminate and release the nascent peptide. After the high salt wash, the stop codon peak in eIF5Ad cells is dramatically higher (~12-fold) than that seen in WT cells (Figure 5C). Importantly, eIF5Ad cells still show increased levels of 3'-UTR ribosomes after clearing by salt washing (Figure 5D), suggesting that the 3'-UTR-localized ribosomes in eIF5Ad cells are actually translating. These observations are consistent with the fact that defective termination can result in ribosome read-through or frame-shifting at stop codons (Dever and Green, 2012; Pande et al., 1995; Stansfield et al., 1995).

We also asked whether eIF5A stimulates peptide release generally or if it has a stronger effect in specific contexts. First, we analyzed the effect of the final amino acid as this has been shown to affect rates of peptide release (Bjornsson et al., 1996; Woolstenhulme et al., 2013). We find that while eIF5A stimulates peptide release in all contexts, termination at certain C-terminal amino acids (e.g. Thr and Val, Figure S5A) is somewhat more affected by eIF5A. We also investigated the role of the nucleotide after the stop codon, a position that has been argued important for eRF1 binding (Brown et al., 2015; Brown et al., 1990; Matheisl et al., 2015), but we find no robust trends (Figure S5B). This finding is in line with our previous observation *in vitro* that the fourth nucleotide in the stop codon has only a minor effect on eRF1-mediated release (Eyler et al., 2013). Taken together, the profiling results point to globally inefficient termination in

eIF5Ad cells, resulting in the accumulation of terminating ribosomes at stop codons and translating ribosomes in the 3'-UTR regions.

eIF5A stimulates peptidyl-tRNA hydrolysis by eRF1 in vitro

Given that termination is inhibited in cells when eIF5A is depleted, we wanted to ask whether eIF5A has a direct effect on the rate of peptidyl-tRNA hydrolysis using our *in vitro* system. We purified elongation complexes with [^{35}S]-Met-Phe-Lys-tRNA^{Lys} in the P site and a stop codon (UAA) in the A site, incubated them with eRF1:eRF3, and followed the rate of hydrolysis of the [^{35}S]-Met-Phe-Lys (MFK) peptide from tRNA^{Lys} over time (Figure 6A). When hypusinated eIF5A was added to the termination reaction with eRF1:eRF3, the rate of peptidyl-tRNA hydrolysis was stimulated 17-fold from a rate of 1.9 min^{-1} to 32.5 min^{-1} (Figure 6B). The hypusine modification of eIF5A is critical for this activity as unmodified eIF5A has only a 4-fold effect on this same reaction. While our earlier work showed a minor effect (1.7-fold) of eIF5A in a peptide release assay (Saini et al., 2009), the dynamic range of our biochemical system has increased over time. Because these experiments were performed with saturating amounts of eIF5A (Figure S6), we conclude that the stimulatory effects that we document are at the level of k_{cat}.

Given that our elongation complexes were purified prior to the peptide hydrolysis assay by pelleting them over a sucrose cushion, we were concerned that the E-site tRNA might dissociate in the cushion, creating a situation where eIF5A binding to the E site might be more important than it is during elongation within cells. To rule out this possibility, we performed the same termination assay in a "one-pot" manner, omitting the

pelleting step, and found that eIF5A still robustly stimulated the termination reaction (Figure S7A and B). These data suggest that the rate of E-site tRNA departure is fast enough in a typical elongation reaction to provide access to the E site for eIF5A. Finally, we performed the termination assay with a catalytically inactive GGQ→AGQ eRF1 to confirm that eIF5A was stimulating hydrolysis through the canonical eRF1-mediated reaction. Indeed, this eRF1 mutant completely abolishes peptidyl-tRNA hydrolysis even in presence of eIF5A (Figure S7C). Taken together, our *in vitro* biochemistry and ribosome profiling provide independent support that eIF5A plays a crucial role in translation termination by stimulating the catalytic activity of eRF1.

eIF5A acts globally while EFP is proline specific

In light of this critical function for eIF5A in translation termination, we wondered if EFP might play a similar role in termination in bacterial translation. We performed the same analysis of stop codon-aligned ribosome footprints with data from *Δefp* cells (Woolstenhulme et al., 2015) and find no evidence that the loss of EFP enhances pausing at stop codons (Figure 7A). Additionally, recent biochemical work using an *E. coli in vitro* reconstituted translation system has shown that EFP does not stimulate the rate of release of fMet-Pro or fMet-Gly from tRNA by RF1 or RF2 (Pierson et al., 2016). These data argue that the function of eIF5A in translation termination is unique to eIF5A.

We also revisited the activities of eIF5A and EFP in promoting peptidyl transfer during elongation by directly comparing their pausing spectra. It is well established that both proteins resolve pauses at polyproline stretches (Doerfel et al., 2013; Elgamal et al., 2014; Gutierrez et al., 2013; Hersch et al., 2013; Peil et al., 2013; Starosta et al., 2014;

Ude et al., 2013; Woolstenhulme et al., 2015). In our eIF5A depletion strain, we identified many tripeptide motifs that pause the ribosome (Table S1), greatly expanding the scope of substrates that eIF5A acts on. Many of these sequences contain no proline at all, but rather combinations of aspartic acid, glycine, and other amino acids. Strong pausing at such motifs was not observed in previous studies of *Δefp E. coli* cells (Elgamal et al., 2014; Woolstenhulme et al., 2015). While the absolute intensity of pauses in these studies cannot be compared directly, differences in the relative enrichment of the pause scores in mutant strains leads to a striking observation. In *Δefp* cells, only the motifs with Pro-Pro lie off the diagonal and are enriched compared to those motifs in wild-type *E. coli* cells (Figure 7B); in contrast, upon eIF5A depletion the entire distribution of pausing motifs is reoriented more vertically, indicative of increased levels of pausing across a majority of tripeptide motifs in the absence of eIF5A (Figure 7C).

This observation explains the defect in global elongation rate upon eIF5A depletion that was previously reported (Saini et al., 2009) and that we observe by ribosome profiling (Figure 1C). In *Δefp* cells we find no evidence of a global redistribution of ribosomes (Figure 7D), but only identify genes that contain the strongest Pro-Pro dipeptide pausing motifs early in the gene to have an effect on the overall distribution of ribosome occupancy (Figure S2C). These comparisons argue that the global effects of eIF5A on translation elongation are unique to eIF5A, while the role of EFP is limited to translation of genes containing proline stretches.

DISCUSSION

eIF5A and EFP were originally identified because they promote formation of (f)Met-Puromycin in a simplified system for studying translation (Benne and Hershey, 1978; Glick et al., 1979; Glick and Ganoza, 1975; Kemper et al., 1976; Schreier et al., 1977) and only later were they shown to promote translation elongation during Pro-Pro bond formation (Doerfel et al., 2013; Gutierrez et al., 2013; Ude et al., 2013). In this study, we report a much broader role for eIF5A in promoting elongation at many (perhaps all) sites in the transcriptome. From ribosome profiling of an eIF5A depletion strain, we identified many tripeptide motifs that substantially pause the ribosome (pause score > 10) (Figure 2 and Table S1) and showed that these widespread pauses lead to a global defect in elongation as observed by ribosome redistribution toward the 5'-end of genes (Figure 1). Our data help to explain prior observations of global defects in elongation (Saini et al., 2009) and support a model where eIF5A is a very general elongation factor. In contrast, EFP is much more specialized, limited to a role in enhancing the rate of translation of Pro-Pro motifs; the loss of EFP does not lead to global defects in elongation (Figure 7).

In this study we also uncovered a role for eIF5A in termination, increasing the rate of eRF1-mediated peptidyl-tRNA hydrolysis. *In vivo*, our ribosome profiling reveals a global defect in termination in the absence of eIF5A (Figure 5), while our previous study of *E. coli* cells lacking EFP shows no such effect (Figure 7). We also showed *in vitro* that eIF5A stimulates eRF1-mediated peptide release by 17-fold (Figure 6) and that this enhanced rate functions through canonical GGQ-hydrolysis (Figure S7C). In contrast to eIF5A, EFP was shown to have no function in RF1 or RF2-mediated peptidyl

hydrolysis (Pierson et al., 2016). These observations support a model where eIF5A functions globally to stimulate slow reactions at the ribosomal peptidyl transferase center (PTC) while EFP is more specialized for several specific slow peptide motifs.

The differences in substrate specificity of these two proteins may be explained at a structural level. In the structure bound to the 70S ribosome, EFP makes extensive contacts with the P-site peptidyl-tRNA (Blaha et al., 2009). These contacts allow EFP to recognize specific substrate tRNAs containing a unique structural element in the D-arm of both $tRNA^{Pro}$ and $tRNA^{fMet}$ (Katoh et al., 2016). In the recent structures of eIF5A bound in the ribosome (Melnikov et al., 2016; Schmidt et al., 2016), direct contact with the D-arm of the P-site tRNA is not observed. Based on the broader role for eIF5A we present here, including biochemical data to support a role for eIF5A function at many peptide sequences (Figures 3 and 4), we do not believe that eIF5A recognizes specific tRNAs as EFP does.

Previous biochemical probing and structural studies have shown that eIF5A binding in the ribosomal E site likely stabilizes the CCA-end of the peptidyl-tRNA (Gutierrez et al., 2013; Melnikov et al., 2016; Schmidt et al., 2016). These observations suggest that when a ribosome encounters problematic motifs (slow peptidyl transfer or eRF1-mediated hydrolysis), the ribosomal E site becomes vacant; prolonged E-site vacancy may allow eIF5A to bind and resolve ribosome pausing. From our data we find eIF5A to alleviate ribosome stalls that occur because of particular amino acid sequences, at the level of promoting slow peptide bond formation (Figures 2 and 3). Interestingly these sequences are not selected against in *S. cerevisiae* genome (data not shown). While other mechanisms of ribosome stalling have been described involving codon optimality

or tRNA identity/abundance, we find no evidence for these types of stalling in our data. For example, we analyzed our profiling data after eIF5A depletion for any stalling effects correlated with codon optimality (Radhakrishnan et al., 2016) and saw no correlation (Figure S3C). Similarly, we found no evidence for stalling at recently identified codon pairs that inhibit translation (Gamble et al., 2016) (Figure S3D; data not shown). Taken together our data suggest that eIF5A is specific for resolving stalls due to slow peptidyl transfer at the ribosomal PTC in the 60S subunit, rather than any events related to the efficiency of tRNA binding or decoding.

Our reported function for eIF5A in termination, in addition to its role in alleviating widespread elongation pausing, may account for both its great abundance and essential nature in eukaryotic cells. In *S. cerevisiae*, eIF5A is one of the most well expressed proteins with > 273,000 copies per cell (Kulak et al., 2014). In fact, eIF5A is in the top 50 highest expressing genes *in S. cerevisiae*, *S. pombe*, and HeLa cells (Kulak et al., 2014), with levels equivalent to that of ribosomes (von der Haar, 2008). In addition to its abundance, eIF5A interacts with 80S ribosomes very strongly, with an approximate dissociation constant of 9 nM (Rossi et al., 2016), suggesting that it could interact with all ribosomes that have an empty E site (eIF5A cellular concentration is approximately 8-15 μM). Given its critical and diverse roles in both translation elongation and termination, these observations lead us to suspect that most, if not all, ribosomes interact with eIF5A throughout the elongation and termination phases of translation.

MATERIALS AND METHODS

Yeast strains and growth conditions

As shown in Figure S1, WT and eIF5Ad strains were grown in YPGR (0.2% galactose+ 0.2% raffinose) overnight at 30°C, harvested by centrifugation, washed and resuspended in YPD (0.2% glucose) medium containing 0.5 mM auxin (Sigma) to an OD_{600} of 0.003 for WT and 0.15 for eIF5Ad. Both WT and eIF5Ad cells were grown in the presence of auxin for 10 hr (to deplete eIF5A completely), harvested by fast filtration, and then flash frozen in liquid nitrogen. It took ~1 min to collect ~800 mL culture. Cell pellets were ground with droplets of footprint lysis buffer [20 mM Tris-Cl (pH8.0), 140 mM KCl, 1.5 mM MgCl2, 1% Triton X-100, 0.1 mg/mL cycloheximide] in a Spex 6870 freezer mill.

Yeast strain construction

ADH1p-OsTIR1 cassette was amplified from pMK200 (Nishimura and Kanemaki, 2014) and inserted in HO locus in BY4741 and Δanb1 strains (GE Dharmacon), resulting in yCW30 (WT strain: MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 HO::ADH1p-OsTIR1-URA3) and yCW31 (MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 anb1::KanMX HO::ADH1p-OsTIR1-URA3)), respectively. After transforming mAID-HYP2-pAG413GAL into yCW31, the genomic copy of *hyp2* was deleted by a NatMX4 cassette, resulting in yCW33 (eIF5Ad strain: MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 HO::adh1p-OsTIR1-URA3 anb1::KanMX hyp2::NatMX4 mAID-HYP2-pAG413GAL).

Preparation of ribosome footprint libraries

After clarification, 25 OD_{260} units of lysates were treated with 375 units of RNaseI (Ambion) for 1 hr at RT. For high salt wash experiments, same amount of lysates were treated with 1 M KCl (final conc.) on ice for 15 min, and desalted by passing through desalting a column (Zeba spin, Thermo Fisher Scientific). The resulting cell lysates were then treated with RNaseI as mentioned above. Monosomes were isolated by sucrose gradients. The extracted RNA was size-selected from 15% denaturing PAGE gels, cutting between 15-34 nt. An oligonucleotide adapter was ligated to the 3' end of footprints. After rRNA depletion using RiboZero (Illumina), reverse transcription, circularization and PCR amplification. Footprint libraries were sequenced on a HiSeq2500 machine at facilities at Johns Hopkins Institute of Genetic Medicine.

Purification of translation factors

Translation initiation factors eIF2, eIF1, eIF1A, eIF5, and eIF5b were purified from both *S. cerevisiae* (eIF2) and *E. coli* using reported procedures (Acker et al., 2007). Purification of eEF1A from *S. cerevisiae* also followed a previously reported protocol (Eyler and Green, 2011).

During purification of eEF2 and eEF3 from *S. cerevisiae* we identified the presence of low-level (though highly active) contaminating eIF5A. Endogenous eIF5A is very abundant, and can be enriched from yeast lysate by both Ni-NTA and amylose affinity chromatography (see Figure S6 for yeast-purified His-MBP activity). In order to isolate eEF2 and eEF3 in the absence of eIF5A, we took great care to sacrifice any contaminated fractions during purification to yield a highly pure protein. After

purification, individual fractions were tested for MPPK synthesis in the absence of eIF5A as a test for contamination. Those fractions that did not stimulate MPPK formation were used in subsequent assays.

eEF2 was purified using a C-terminal His_6 tag from strain TKY675 (Jorgensen et al., 2002). eEF2 was purified by Ni-NTA chromatography in buffer containing 20 mM HEPES pH 7.4, 10% glycerol, 300 mM KCl, and 5 mM 2-mercaptoethanol. Then the protein was diluted to 30 mM KCl and further purified by cation exchange chromatography using a gradient from 30-150 mM KCl. Fractions containing eEF2 (and no eIF5A) were pooled, concentrated, and further purified by gel filtration using a HiPrep 26/60 Sephacryl S-200 column (GE Healthcare Life Sciences) for best separation of eEF2 and contaminant eIF5A. Fractions containing eEF2 were individually concentrated, and stored in gel filtration buffer containing 20 mM HEPES pH 7.4, 100 mM KOAc pH 7.6, 10% glycerol, and 2 mM DTT.

eEF3 was purified using a N-terminal His_6 tag from strain TKY702 (Andersen et al., 2004). eEF3 was purified by Ni-NTA chromatography in buffer containing 25 mM Tris pH 7.5, 10 mM Imidazole, 1 M KCl, 10% glycerol, and 5 mM 2-mercaptoethanol. After elution, protein was further purified by gel filtration using a Superose 6 10/300GL column (GE Healthcare Life Sciences). Fractions containing eEF3 were individually concentrated, and stored in gel filtration buffer containing 20 mM HEPES pH 7.4, 100 mM KOAc pH 7.6, 10% glycerol, and 2 mM DTT.

Eukaryotic release factor eRF1 (and AGQ mutant) was purified from *E. coli* with a His_6 tag as previously described (Shoemaker et al., 2010). Release factor eRF3 was

purified from *E. coli* using the IMPACT Protein Purification System (New England Biolabs) as previously described (Eyler et al., 2013).

Expression and purification of eIF5A

Expression and purification of recombinant eIF5A from *E. coli* was performed as previously reported (Gutierrez et al., 2013). Briefly, a plasmid expressing His_6-eIF5A, or a plasmid co-expressing His_6-eIF5A with modification enzymes Dys1 and Lia1 was transformed into BL21-CodonPlus (DE3)-RIPL cells (Agilent). Proteins were expressed and purified by Ni-NTA affinity chromatography and anion exchange chromatography.

To confirm eIF5A hypusination, proteins were analyzed by MALDI-TOF mass spectrometry (Voyager DE-STR MALDI-TOF). As shown in Figure S4B, unmodified eIF5A purified from *E. coli* showed a major peak with a calculated weight of 17804, consistent with the predicted mass of unmodified His_6-eIF5A. Analysis of His_6-eIF5A purified from *E. coli* expressing Dys1 and Lia1 modification enzymes revealed a shift in molecular weight to 17901, consistent with modification.

Purification of tRNAs and mRNAs

Phenylalanine-specific tRNA from *S. cerevisiae* was purchased from Sigma. Initiator methionine, lysine and cysteine-specific tRNAs were purchased from tRNA Probes (College Station, TX). Proline, aspartic acid, and isoleucine tRNAs were isolated from *S. cerevisiae* bulk tRNA using 3' biotinylated oligos (IDT) as previously described (Yokogawa et al., 2010). Oligo for $tRNA^{Pro}$(UGG): 5' – CCAAAGCGAG AATCATACCA CTAGAC – 3' Biotin-TEG. Oligo for $tRNA^{Asp}$(GUC): 5' –

GACAAGCGCCCATTCTGACCATTAAAC – 3' Biotin-TEG. Oligo for $tRNA^{Ile}$(AAU): 5' – ATTAGCACGGTGCCTTAACCAACTGGGC – 3' Biotin-TEG. Isolated $tRNA^{Pro}$, $tRNA^{Asp}$, and $tRNA^{Ile}$ were subject to CCA-adding reaction (Gutierrez et al., 2013). All tRNAs were charged using a S100 extract (Eyler and Green, 2011).

Model mRNAs for *in vitro* translation were transcribed using T7 RNA polymerase and contain the following sequence: 5' – GAAUCUCUCUCUCUCUCU **AUG XXX XXX XXX UAA** CUCUCUCUCUCUCUC – 3' (underlined G is T7 start, bold is open reading frame, XXX = generic codon). In elongation experiments, codons used are as follows: Phe (UUC), Pro (CCA), Asp (GAC), Ile (AUU), Cys (UGC), and Lys (AAA). All mRNAs were gel purified using 10% TBE-Urea gels.

In vitro 80S initiation and elongation complex formation

80S initiation complexes were formed as previously described (Eyler and Green, 2011) with minor differences. Briefly, 3 pmol of ^{35}S-Met-$tRNA^{iMet}$ was mixed with 25 pmol of eIF2 and 1 mM GTP in 1X Buffer E (20 mM Tris pH 7.5, 100 mM KOAc pH 7.6, 2.5 mM $Mg(OAc)_2$, 0.25 mM Spermidine, and 2 mM DTT) for 10 minutes at 26^{o}C. Next a mixture containing 25 pmol 40S subunits, 200 pmol T7-synthesized mRNA template, 150 pmol eIF1, and 150 pmol eIF1A in 1X Buffer E was added for 5 minutes. To form the 80S complex, a mixture containing 25 pmol 60S subunits, 125 pmol eIF5, 125 pmol eIF5b, and 1 mM GTP in 1X Buffer E was added for 1 minute. Complexes were then mixed 1:1 with buffer E containing 17.5 mM $Mg(OAc)_2$ to yield a final magnesium concentration of 10 mM. Ribosomes were then pelleted through a 600 µl sucrose cushion containing 1.1 M sucrose in buffer E with 10 mM $Mg(OAc)_2$ using a

MLA-130 rotor (Beckmann) at 75,000 rpm for 1 hour at 4°C. After pelleting, ribosomes were resuspended in 25 µl of 1X Buffer E containing 10 mM $Mg(OAc)_2$ and stored at -80°C.

To form elongation complexes for peptide release experiments, initiation complexes were formed as described above and elongated before pelleting. To create the eEF1A ternary complex, a mixture containing 50 pmol eEF1A, 30 pmol aa-tRNA, and 1mM GTP was incubated at 26°C for 15 minutes. Ternary complex for each required tRNA was mixed with 80S initiation complexes for 5 minutes to allow for peptidyl transfer. Subsequently a mixture containing 50 pmol eEF2, 75 pmol eEF3, 2.5 mM GTP, and 5.0 mM ATP was added to promote elongation and subsequent rounds of peptidyl transfer for 7 minutes. Elongated complexes were then mixed 1:1 with buffer E containing 17.5 mM $Mg(OAc)_2$ and pelleted as described above.

***In vitro* reconstituted translation elongation**

Translation elongation reactions were performed in 1X Buffer E (20 mM Tris pH 7.5, 100 mM KOAc pH 7.6, 2.5 mM $Mg(OAc)_2$, 0.25 mM Spermidine, and 2 mM DTT). Limited amounts of 80S initiation complexes (3 nM) were mixed with purified eEF1A (1 µM), aa-tRNA (500 nM), eEF2 (1 µM), eEF3 (1 µM), ATP (3 mM) and GTP (2 mM) in the presence or absence of eIF5A (1 µM). Reactions were incubated at 26°C and time points quenched with 250 mM KOH. Peptide formation was monitored by electrophoretic TLC (Millipore). TLC plates were equilibrated with pyridine acetate buffer (5 ml pyridine, 200 ml acetic acid in 1 liter, pH 2.8) before electrophoresis at 1200 V for 28 min. Plates were developed using a Typhoon FLA 9500 Phosphorimager system (GE

Healthcare Life Sciences) and quantified using ImageQuantTL (GE Healthcare Life Sciences). Time courses were fit to single exponential kinetics using Kaleidagraph (Synergy Software).

In vitro reconstituted eRF1:eRF3 peptidyl hydrolysis

Peptide release assays were performed in 1X Buffer E (same as elongation reactions). Limited amounts of 80S elongation complexes (3 nM) were mixed with purified eRF1 (4 μM), eRF3 (6 μM) and GTP (1 mM) in the presence or absence of eIF5A (1 μM). Reactions were incubated at 26^{o}C and time points quenched with 5% formic acid. TLC plates were equilibrated with pyridine acetate buffer (5 ml pyridine, 200 ml acetic acid in 1 liter, pH 2.8) before electrophoresis at 1200 V for 18 min. Plates were developed, quantified, and fit to single exponential kinetics as described for elongation reactions.

In vitro Met-Puromycin assay

Reactions containing 2 nM initiation complexes and 1 μM eIF5A in 1X Buffer E (20 mM Tris pH 7.5, 100 mM KOAc pH 7.6, 2.5 mM $Mg(OAc)_2$, 0.25 mM Spermidine, and 2 mM DTT) were incubated at 26^{o}C in the presence of 5 mM puromycin. Time points over the course of 90 minutes were quenched with 250 mM KOH and analyzed by electrophoretic TLC (Millipore). TLC plates were equilibrated with pyridine acetate buffer (5 ml pyridine, 200 ml acetic acid in 1 liter, pH 2.8) before electrophoresis at 1200 V for 15min. Plates were developed using a Typhoon FLA 9500 Phosphorimager system (GE Healthcare Life Sciences) and quantified using ImageQuantTL (GE Healthcare Life

Sciences). Time courses were fit to single exponential kinetics using Kaleidagraph (Synergy Software).

In vitro analysis of peptidyl-tRNA dropoff using PTH

Translation elongation reactions were performed as described in the primary methods in the presence of 27 μM peptidyl-tRNA hydrolase (PTH) to monitor dropoff of peptidyl-tRNAs from translating ribosomes. PTH can only cleave peptidyl-tRNAs that have been fully released from the ribosome (Figure S4D, note zero time point). Time points for dropoff products were quenched with 10% formic acid, and peptide formation was monitored by KOH quench at the start and end of the reaction. Reaction time points were analyzed by electrophoretic TLC in pyridine acetate buffer (see above) at 1200 V for 30 minutes.

Read preparation and alignment

The R64-1-1 S288C reference genome assembly (SacCer3) from the *Saccharomyces* Genome Database Project was used for all analyses. Adapter sequence (CTGTAGGCACCATCAAT) was first removed from demultiplexed reads using Cutadapt (Martin, 2011), and low quality reads (any position with Phred score less than 20) were discarded. Then alignment to the RNA gene database FASTA file (http://downloads.yeastgenome.org/sequence/S288C_reference/rna/archive/rna_coding_R64-1-1_20110203.fasta.gz) was performed to remove noncoding RNAs. The resulting reads were then aligned to the genome using 3' end mapping, and the reads that failed to be mapped were aligned to the annotated splice junctions Finally, reads left were trimmed

of consecutive As from their 3' ends and realigned to the genome and the splice junction. Read length between 25-34 nt was used for the following analyses. The mapped reads were normalized to reads per million (rpm) using total number of mapped reads. Alignment was performed using Bowtie 1.0.0 (Langmead et al., 2009) using the following parameters: '-v 2 –y –a –m 1 –best –strata –S –p 4'. All other analyses were performed using software custom written in Python 2.7 and R 3.3.1.

Ribosome profiling datasets for *Δefp* (Woolstenhulme et al., 2015) were downloaded from the GEO (GSE64488) and the *Escherichia coli* MG1655 reference genome NC_000913.2 was used for all *E. coli* ribosome profiling analyses.

Analysis of aligned reads

Calibrated by using start and stop codons of coding sequences, we defined that the first 5' end nucleotide of ribosome A site to the 3' end of the footprint is separated by 12 nt. In general, -14 shift (the center of P site) is used for start codons (Figure 1A), pause motifs and quantitation of coding sequences (Figures 1C, 1D, 2, and 7), and -11 shift (center of ribosomal A site) is used for stop codons (Figure 5).

The polarity at position *i* in a gene of length *l* is defined as follows:

$$p_i = \frac{d_i w_i}{\sum_{i=1}^{l} d_i}$$

where

$$w_i = \frac{2i-(l+1)}{l-1}$$

The terms d_i and w_i are the ribosome density and the normalized distance from the center of a gene at position *i*, respectively. Polarity score for a gene is the total sum of p_i at each position. To avoid known artifacts around start and stop codon peaks, we excluded the

first and the last 15 nt of coding sequences from our analysis. Genes with more than 64 reads/dataset in coding sequences were plotted.

Pause scores for peptide motifs (Figures 2D, 7B and 7C) were calculated by taking the ratio of the ribosome density in a 3-nt window at the motif over the overall density in the coding sequence (excluding the first and the last 100 nt). Genes with less than 64 reads/dataset in coding sequences (excluding the first and the last 100 nt) are excluded from the analysis. The same threshold was used for metagene plots (Figures 1A and 5) and metacodon plots (Figures 2A, 2B and 2E). Logo was weighted by pause score ratio of eIF5Ad/WT (Table S1). For metagene plots, ribosome density at each position is normalized by the overall ribosome density of coding sequence. Genes with features that are smaller than the window size (50 nt upstream of start codons and 900 nt of coding sequences for Figure 1A, and 100 nt upstream of stop codons and 50 nt of 3'-UTRs for Figure 5) were excluded from the analysis.

Accession numbers:

Sequencing data were deposited in the GEO database under the accession number GSE89704.

TLC and gel images are deposited in the Mendeley database at the following link: http://dx.doi.org/10.17632/nmkd3jbhx7.1

FIGURE LEGENDS

Figure 1. Polar distribution of ribosomes toward 5' end of genes in eIF5Ad cells

(A) Average ribosome occupancy from all genes aligned at start codons for WT (black line) and eIF5Ad cells (orange line) with schematic depicting queuing of ribosomes upstream of a paused ribosome (colored orange). Ribosome occupancy was normalized to show a mean value of 1 for each codon. (B) Example of ribosome occupancy along *CHC1* gene in WT.1 (black) and eIF5Ad.1 (orange) cells. Rpm, reads per million. (C) Distributions of polarity scores for 4,946 genes from WT and eIF5Ad cells are plotted, for genes with at least 64 reads/dataset in ORFs (top panel). Schematic representation of polarity score (bottom panel). (D) Distributions of polarity scores for genes containing Pro-Pro dipeptide motifs (green) and genes lacking these motifs (brown) in eIF5Ad cells. WT distribution (includes both PP and nonPP) is included for reference (dotted). Numbers in parentheses denote the gene numbers included in the analysis. See also Figure S1 and S2.

Figure 2. eIF5A alleviates ribosome pausing at more than poly-Pro motifs

(A) Average ribosome occupancy centered at diproline motifs with the underlined Pro in the P site of the ribosome. Excluding motifs in genes with less than 64 reads/dataset in ORFs, 5,920 and 5,939 positions are averaged in WT and eIF5Ad, respectively. Arrow indicates stacked ribosomes. (B) Similar to (A), average plot centered at triproline motifs with underlined Pro in the E site of the ribosome. Inset: close-up view of the ribosome occupancy 10-50 nt downstream of the triproline motif. (C) Ribosome footprints on *THO1* and *PCL6* genes in WT.1 (black) and eIF5Ad.1 (orange) cells. Motifs of highest

ribosome pausing are denoted. Rpm, reads per million. (D) Peptide motif associated with ribosome pausing in eIF5Ad cells, using motifs with pause score greater than 10 and weighted by pause score ratio of eIF5Ad/WT. Top 29 motifs are listed. (E) Average ribosome occupancy centered at 5,478 pause sites that match the 18 non Pro-Pro pausing motifs identified in eIF5Ad cells. See also Figure S3.

Figure 3. eIF5A stimulates translation of stalling motifs *in vitro*

(A) Schematic representation of *in vitro* elongation reactions using reconstituted translation system. Pelleted 80S initiation complexes are incubated with elongation factors, aminoacyl-tRNAs, GTP, ATP, and eIF5A. Reactions are quenched with KOH and peptide products resolved by electrophoretic TLCs. (B) Elongation kinetics for Phe-Phe and Pro-Pro containing peptides (MFFK and MPPK) in the presence and absence of eIF5A and hypusination modification. (C) Representative TLCs for elongation kinetics of dipeptide stalling motifs Asp-Asp (MDDK) and Asp-Pro (MDPK). (D) Comparison of elongation endpoints for all peptides analyzed in presence and absence of eIF5A. Error bars represent the standard deviation from three replicate experiments. See also Figure S4.

Figure 4. Hypusine is required for polyproline elongation *in vitro*

Translation elongation kinetics for the following tripeptide motifs in the presence and absence of eIF5A and hypusination modification: (A) Phe-Phe-Phe (MFFFK), (B) Pro-Pro-Pro (MPPPK), (C) Pro-Asp-Ile (MPDIK), (D) Asp-Asp-Ile (MDDIK). Time points

were quantified and reaction progression fitted to a single exponential. See also Figure S4.

Figure 5. Ribosomes accumulate at stop codons and in 3'-UTRs in eIF5Ad cells

(A) Metagene analysis of translation termination. Average ribosome occupancy from all genes aligned at their stop codons for WT and eIF5Ad cells. Arrow denotes stacked ribosomes ~30nt upstream the stop codon peak. (B) Overlay and close-up view of (A), showing accumulated ribosomes in 3'-UTRs in eIF5Ad cells. (C) Similar to (A), metagene plot of stop codons for WT and eIF5Ad lysates treated with high salt buffer. (D) Similar to (B), close-up view of 3'-UTR regions for WT and eIF5Ad lysates after high salt wash. See also Figure S5.

Figure 6. eIF5A stimulates eRF1-mediated peptidyl hydrolysis

(A) Schematic for *in vitro* termination assays. Met-Phe-Lys elongation complexes containing an A-site stop codon (UAA) are reacted with eRF1:eRF3:GTP in the presence and absence of eIF5A. Time points are quenched with formic acid, peptide products resolved by electrophoretic TLCs, and rates quantified. (B) Rates of peptidyl hydrolysis by eRF1:eRF3 in presence and absence of eIF5A and hypusination modification. Error bars represent the standard deviation from at least three replicate experiments. See also Figures S6 and S7.

Figure 7. eIF5A has a more general role than EFP

(A) Metagene analysis of translation termination for wt and *Δefp E. coli* cells shows neither stacked ribosomes (~30 nt upstream of stop codons) nor 3'-UTR ribosomes.

(B) Comparison of tripeptide pausing in wt and *Δefp E. coli* cells. Pause scores of tripeptide motifs are plotted using *E. coli* wt and *Δefp* datasets (Woolstenhulme et al., 2015). Motifs with less than 30 occurrences in *E. coli* transcriptome are excluded from the analysis. Each dot represents one tripeptide motif; 6,018 motifs are included. Motifs with pause score higher than 8 are labeled. The diagonal line indicates the distribution expected for no enrichment. (C) Comparison of tripeptide pausing in WT and eIF5Ad cells. Pause scores of 6,022 tripeptide motifs (Table S1) are plotted for WT and eIF5Ad cells, for motifs with more than 100 occurrences in yeast transcriptome. 10 tripeptide motifs with highest pause scores upon eIF5A depletion are labeled. The diagonal line indicates the distribution expected for no enrichment. (D) Distributions of polarity scores for 2,186 genes from wt and *Δefp E. coli* cells are plotted, showing no significant difference. See also Figures S2 and S3.

SUPPLEMENTARY FIGURES AND TABLES

Figure S1. Growth of WT and eIF5Ad strains used for ribosome profiling –Related to Figure 1

(A) Western blot analysis of eIF5A depletion over time. Same number of cells were pelleted, lysed and subjected to western blotting using antibodies against eIF5A (Saini et al., 2009), mAID (MBL International Corporation) and PGK1 (Life Technologies-Novex). (B) Growth of WT and eIF5Ad cells on YPGR and YPDauxin plates. Plates are incubated at 30°C for 2 days. (C) Growth potential of WT and eIF5Ad strains after media change to YPD containing 0.5 mM auxin (0 hr time point) by monitoring growth at OD_{600}. (D) Sucrose gradient sedimentation analysis of cell lysates from WT and eIF5Ad cells. Cells were grown and harvested under the condition in (C), lysed in footprint lysis buffer and subjected to centrifugation through 10-50% sucrose gradients. Polysome profiles were collected by monitoring at 254 nm. P/M denotes polysome/monosome ratio. (E) Scatter plots showing reproducibility of ribosome footprints on coding sequences between biological replicates for WT (left) and eIF5Ad (right). Pearson correlations are shown in each comparison.

Figure S2. Disproportionate distribution of ribosome occupancy in eFI5Ad cells – Related to Figure 1 and 7

(A) Polarity scores of ribosome occupancy on 4,946 genes from Figure 1C are plotted for WT and eIF5Ad cells. Genes in lower left quadrant correspond to those with disproportionate distribution of ribosome occupancy toward their 5'-ends. (B) In comparison with Figure 1D, distributions of polarity scores for genes containing Pro-Pro

dipeptide motifs and genes lacking these motifs from WT cells. (C) Polarity scores of 2,186 genes are plotted for wt and *Δefp E. coli* strains (Woolstenhulme et al., 2015). Genes with strongest Pro-containing stalling motifs before the halfway point are shown in pink.

Figure S3. Metacodon analysis of ribosome stalling in eIF5Ad cells – Related to Figure 2 and 7

(A) Average ribosome occupancy centered at 9,082 pause sites that match the 29 pausing motifs identified in eIF5Ad cells (Figure 2D). Arrow indicates stacked ribosomes. (B) Similar to (A), average plot centered at 1,350 pause sites that match the 4 non-Pro tripeptide motifs (RDK, DVG, DDG, GGT). (C) Analysis of codon optimality effects in eIF5Ad strain. Ribosome occupancy in the E, P and A sites under eIF5A depletion relative to WT. 61 sense codons are colored according to their stAI values (Sabi and Tuller, 2014). (D) Average ribosome density plot centered at the CTCCCG inhibitory codon pair (Gamble et al., 2016).

Figure S4. eIF5A purification and modification, stimulation of Met-Puromycin formation, and analysis of *in vitro* peptidyl-tRNA dropoff – Related to Figures 3 and 4

(A) Coomassie blue stained gel of purified eIF5A proteins. (B) MALDI-TOF analysis of purified eIF5A proteins to confirm hypusine modification. (C) Electrophoretic TLC analysis of *in vitro* Met-Puromycin assay. Time points of Met-Puro formation were quantified and reaction progression fitted to a single exponential to calculate the observed

rate. (D) Electrophoretic TLC analysis of dropoff products in elongation reactions as observed by addition of peptidyl-tRNA hydrolase (PTH).

Figure S5. Effect of C-terminal amino acid and +4 nucleotide on eIF5A-stimulated termination – Related to Figure 5

(A) Violin plot of pause scores at stop codons in WT and eIF5Ad cells sorted by C-terminal amino acid. Fold change indicates the average pause score ratio of eIF5Ad/WT. (B) Similar to (A), but sorted by the nucleotide following the stop codon.

Figure S6. eIF5A is a potent stimulator of peptide release activity, is heat-insensitive, and is a common contaminant of protein purifications – Related to Figure 6

(A) Electrophoretic TLC analysis of eRF1:eRF3 peptide release kinetics in the presence of various amounts of hyspusinated eIF5A, hypusinated eIF5A that was boiled at 95^{o}C for 15 min prior to addition, unmodified eIF5A, and a purified His_6-MBP construct expressed in *S. cerevisiae* containing low-level (though highly active) eIF5A contamination (see protein purification methods for more discussion). (B) Observed rates of peptide hydrolysis calculated for samples presented in (A).

Figure S7. eIF5A stimulates peptide release on pelleted elongation complexes, non-pelleted complexes, and functions through canonical eRF1-GGQ mechanism – Related to Figure 6

(A) Electrophoretic TLC analysis of eRF1:eRF3 peptide release kinetics in the presence of hypusinated eIF5A using pelleted MFK complexes containing UAA stop codon in A site. (B) Same as (A) except using elongation complexes that were not pelleted, containing MFFK peptide and UAA stop codon in A site. TLCs were cropped to include similar time points for comparison. (C) Analysis of eIF5A stimulation of peptide release kinetics in presence of wild-type eRF1 and an eRF1 GGQ→AGQ mutant that is defective for peptidyl hydrolysis.

Table S1. Pausing at tripeptide motifs in WT and eIF5Ad cells – Related to Figure 2 and 7

Pause scores of 6,022 tripeptide motifs are computed for WT and eIF5Ad cells, and ranked by the pause scores in eIF5Ad cells. Motifs with less than 100 occurrences in yeast transcriptome were excluded. Ratio denotes pause score ratio of eIF5Ad/WT. Counts are the numbers of motifs included for average.

Table S2. Pause scores for dipeptide motifs in WT and eIF5Ad cells – Related to Figure 3

Pause scores of 396 dipeptide motifs, ranked by the pause scores in eIF5Ad cells. Motifs with less than 500 occurrences in yeast transcriptome were excluded.

Table S3. Oligonucleotides used in this study

REFERENCES

Acker, M.G., Kolitz, S.E., Mitchell, S.F., Nanda, J.S., and Lorsch, J.R. (2007). Reconstitution of yeast translation initiation. Methods Enzymol *430*, 111-145.

Andersen, C.F., Anand, M., Boesen, T., Van, L.B., Kinzy, T.G., and Andersen, G.R. (2004). Purification and crystallization of the yeast translation elongation factor eEF3. Acta Crystallogr D Biol Crystallogr *60*, 1304-1307.

Artieri, C.G., and Fraser, H.B. (2014). Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. Genome Res *24*, 2011-2021.

Benne, R., and Hershey, J.W. (1978). The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. J Biol Chem *253*, 3078-3087.

Bjornsson, A., Mottagui-Tabar, S., and Isaksson, L.A. (1996). Structure of the C-terminal end of the nascent peptide influences translation termination. EMBO J *15*, 1696-1704.

Blaha, G., Stanley, R.E., and Steitz, T.A. (2009). Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science *325*, 966-970.

Blobel, G., and Sabatini, D. (1971). Dissociation of mammalian polyribosomes into subunits by puromycin. Proc Natl Acad Sci U S A *68*, 390-394.

Brown, A., Shao, S., Murray, J., Hegde, R.S., and Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. Nature *524*, 493-496.

Brown, C.M., Stockwell, P.A., Trotman, C.N., and Tate, W.P. (1990). Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. Nucleic Acids Res *18*, 6339-6345.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol *4*, a013706.

Dever, T.E., Gutierrez, E., and Shin, B.S. (2014). The hypusine-containing translation factor eIF5A. Crit Rev Biochem Mol Biol *49*, 413-425.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85-88.

Elgamal, S., Katz, A., Hersch, S.J., Newsom, D., White, P., Navarre, W.W., and Ibba, M. (2014). EF-P dependent pauses integrate proximal and distal signals during translation. PLoS Genet *10*, e1004553.

Eyler, D.E., and Green, R. (2011). Distinct response of yeast ribosomes to a miscoding event during translation. RNA *17*, 925-932.

Eyler, D.E., Wehner, K.A., and Green, R. (2013). Eukaryotic release factor 3 is required for multiple turnovers of peptide release catalysis by eukaryotic release factor 1. J Biol Chem *288*, 29530-29538.

Gamble, C.E., Brule, C.E., Dean, K.M., Fields, S., and Grayhack, E.J. (2016). Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. Cell *166*, 679-690.

Gerashchenko, M.V., and Gladyshev, V.N. (2014). Translation inhibitors cause abnormalities in ribosome profiling experiments. Nucleic Acids Res *42*, e134.

Glick, B.R., Chladek, S., and Ganoza, M.C. (1979). Peptide bond formation stimulated by protein synthesis factor EF-P depends on the aminoacyl moiety of the acceptor. Eur J Biochem *97*, 23-28.

Glick, B.R., and Ganoza, M.C. (1975). Identification of a soluble protein that stimulates peptide bond synthesis. Proc Natl Acad Sci U S A *72*, 4257-4260.

Gregio, A.P., Cano, V.P., Avaca, J.S., Valentini, S.R., and Zanelli, C.F. (2009). eIF5A has a function in the elongation step of translation in yeast. Biochem Biophys Res Commun *380*, 785-790.

Grousl, T., Ivanov, P., Malcova, I., Pompach, P., Frydlova, I., Slaba, R., Senohrabkova, L., Novakova, L., and Hasek, J. (2013). Heat shock-induced accumulation of translation elongation and termination factors precedes assembly of stress granules in S. cerevisiae. PLoS One *8*, e57083.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Mol Cell *51*, 35-45.

Guydosh, N.R., and Green, R. (2014). Dom34 rescues ribosomes in 3' untranslated regions. Cell *156*, 950-962.

Henderson, A., and Hershey, J.W. (2011). Eukaryotic translation initiation factor (eIF) 5A stimulates protein synthesis in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *108*, 6415-6419.

Hersch, S.J., Wang, M., Zou, S.B., Moon, K.M., Foster, L.J., Ibba, M., and Navarre, W.W. (2013). Divergent protein motifs direct elongation factor P-mediated translational regulation in Salmonella enterica and Escherichia coli. MBio *4*, e00180-00113.

Hussmann, J.A., Patchett, S., Johnson, A., Sawyer, S., and Press, W.H. (2015). Understanding Biases in Ribosome Profiling Experiments Reveals Signatures of Translation Dynamics in Yeast. PLoS Genet *11*, e1005732.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218-223.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Jorgensen, R., Carr-Schmid, A., Ortiz, P.A., Kinzy, T.G., and Andersen, G.R. (2002). Purification and crystallization of the yeast elongation factor eEF2. Acta Crystallogr D Biol Crystallogr *58*, 712-715.

Katoh, T., Wohlgemuth, I., Nagano, M., Rodnina, M.V., and Suga, H. (2016). Essential structural elements in tRNA(Pro) for EF-P-mediated alleviation of translation stalling. Nat Commun *7*, 11657.

Kemper, W.M., Berry, K.W., and Merrick, W.C. (1976). Purification and properties of rabbit reticulocyte protein synthesis initiation factors M2Balpha and M2Bbeta. J Biol Chem *251*, 5551-5557.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Methods *11*, 319-324.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal *17*, 10-12.

Matheisl, S., Berninghausen, O., Becker, T., and Beckmann, R. (2015). Structure of a human translation termination complex. Nucleic Acids Res *43*, 8615-8626.

Melnikov, S., Mailliot, J., Shin, B.S., Rigger, L., Yusupova, G., Micura, R., Dever, T.E., and Yusupov, M. (2016). Crystal Structure of Hypusine-Containing Translation Factor eIF5A Bound to a Rotated Eukaryotic Ribosome. J Mol Biol *428*, 3570-3576.

Mills, E.W., Wangen, J., Green, R., and Ingolia, N.T. (2016). Dynamic Regulation of a Ribosome Rescue Pathway in Erythroid Cells and Platelets. Cell Rep *17*, 1-10.

Nishimura, K., and Kanemaki, M.T. (2014). Rapid Depletion of Budding Yeast Proteins via the Fusion of an Auxin-Inducible Degron (AID). Curr Protoc Cell Biol *64*, 20 29 21-16.

Pande, S., Vimaladithan, A., Zhao, H., and Farabaugh, P.J. (1995). Pulling the ribosome out of frame by +1 at a programmed frameshift site by cognate binding of aminoacyl-tRNA. Mol Cell Biol *15*, 298-304.

Park, J.H., Dias, C.A., Lee, S.B., Valentini, S.R., Sokabe, M., Fraser, C.S., and Park, M.H. (2011). Production of active recombinant eIF5A: reconstitution in E.coli of

eukaryotic hypusine modification of eIF5A by its coexpression with modifying enzymes. Protein Eng Des Sel *24*, 301-309.

Park, M.H. (1989). The essential role of hypusine in eukaryotic translation initiation factor 4D (eIF-4D). Purification of eIF-4D and its precursors and comparison of their activities. J Biol Chem *264*, 18531-18535.

Park, M.H., Cooper, H.L., and Folk, J.E. (1981). Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. Proc Natl Acad Sci U S A *78*, 2869-2873.

Park, M.H., Wolff, E.C., Smit-McBride, Z., Hershey, J.W., and Folk, J.E. (1991). Comparison of the activities of variant forms of eIF-4D. The requirement for hypusine or deoxyhypusine. J Biol Chem *266*, 7988-7994.

Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M., and Forster, A.C. (2009). Slow peptide bond formation by proline and other N-alkylamino acids in translation. Proc Natl Acad Sci U S A *106*, 50-54.

Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumae, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., and Wilson, D.N. (2013). Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. Proc Natl Acad Sci U S A *110*, 15265-15270.

Pierson, W.E., Hoffer, E.D., Keedy, H.E., Simms, C.L., Dunham, C.M., and Zaher, H.S. (2016). Uniformity of Peptide Release Is Maintained by Methylation of Release Factors. Cell Rep *17*, 11-18.

Radhakrishnan, A., Chen, Y.H., Martin, S., Alhusaini, N., Green, R., and Coller, J. (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell *167*, 122-132 e129.

Rossi, D., Barbosa, N.M., Galvao, F.C., Boldrin, P.E., Hershey, J.W., Zanelli, C.F., Fraser, C.S., and Valentini, S.R. (2016). Evidence for a Negative Cooperativity between eIF5A and eEF2 on Binding to the Ribosome. PLoS One *11*, e0154205.

Sabi, R., and Tuller, T. (2014). Modelling the efficiency of codon-tRNA interactions based on codon usage bias. DNA Res *21*, 511-526.

Saini, P., Eyler, D.E., Green, R., and Dever, T.E. (2009). Hypusine-containing protein eIF5A promotes translation elongation. Nature *459*, 118-121.

Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2016). Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. Nucleic Acids Res *44*, 1944-1951.

Schreier, M.H., Erni, B., and Staehelin, T. (1977). Initiation of mammalian protein synthesis. I. Purification and characterization of seven initiation factors. J Mol Biol *116*, 727-753.

Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. Science *330*, 369-372.

Stansfield, I., Akhmaloka, and Tuite, M.F. (1995). A mutant allele of the SUP45 (SAL4) gene of Saccharomyces cerevisiae shows temperature-dependent allosuppressor and omnipotent suppressor phenotypes. Curr Genet *27*, 417-426.

Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Woolstenhulme, C.J., Virumae, K., Buskirk, A., Tenson, T., Remme, J., Jung, K.*, et al.* (2014). A conserved proline triplet in Val-tRNA synthetase and the origin of elongation factor P. Cell Rep *9*, 476-483.

Subramaniam, A.R., Zid, B.M., and O'Shea, E.K. (2014). An integrated approach reveals regulatory controls on bacterial translation elongation. Cell *159*, 1200-1211.

Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., and Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science *339*, 82-85.

von der Haar, T. (2008). A quantitative estimation of the global translational activity in logarithmically growing yeast cells. BMC Syst Biol *2*, 87.

Weinberg, D.E., Shah, P., Eichhorn, S.W., Hussmann, J.A., Plotkin, J.B., and Bartel, D.P. (2016). Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. Cell Rep *14*, 1787-1799.

Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M.V. (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. J Biol Chem *283*, 32229-32235.

Woolstenhulme, C.J., Guydosh, N.R., Green, R., and Buskirk, A.R. (2015). High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. Cell Rep *11*, 13-21.

Woolstenhulme, C.J., Parajuli, S., Healey, D.W., Valverde, D.P., Petersen, E.N., Starosta, A.L., Guydosh, N.R., Johnson, W.E., Wilson, D.N., and Buskirk, A.R. (2013). Nascent peptides that block protein synthesis in bacteria. Proc Natl Acad Sci U S A *110*, E878-887.

Yokogawa, T., Kitamura, Y., Nakamura, D., Ohno, S., and Nishikawa, K. (2010). Optimization of the hybridization-based method for purification of thermostable tRNAs in the presence of tetraalkylammonium salts. Nucleic Acids Res *38*, e89.

Young, D.J., Guydosh, N.R., Zhang, F., Hinnebusch, A.G., and Green, R. (2015). Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation Reinitiation in 3'UTRs In Vivo. Cell *162*, 872-884.

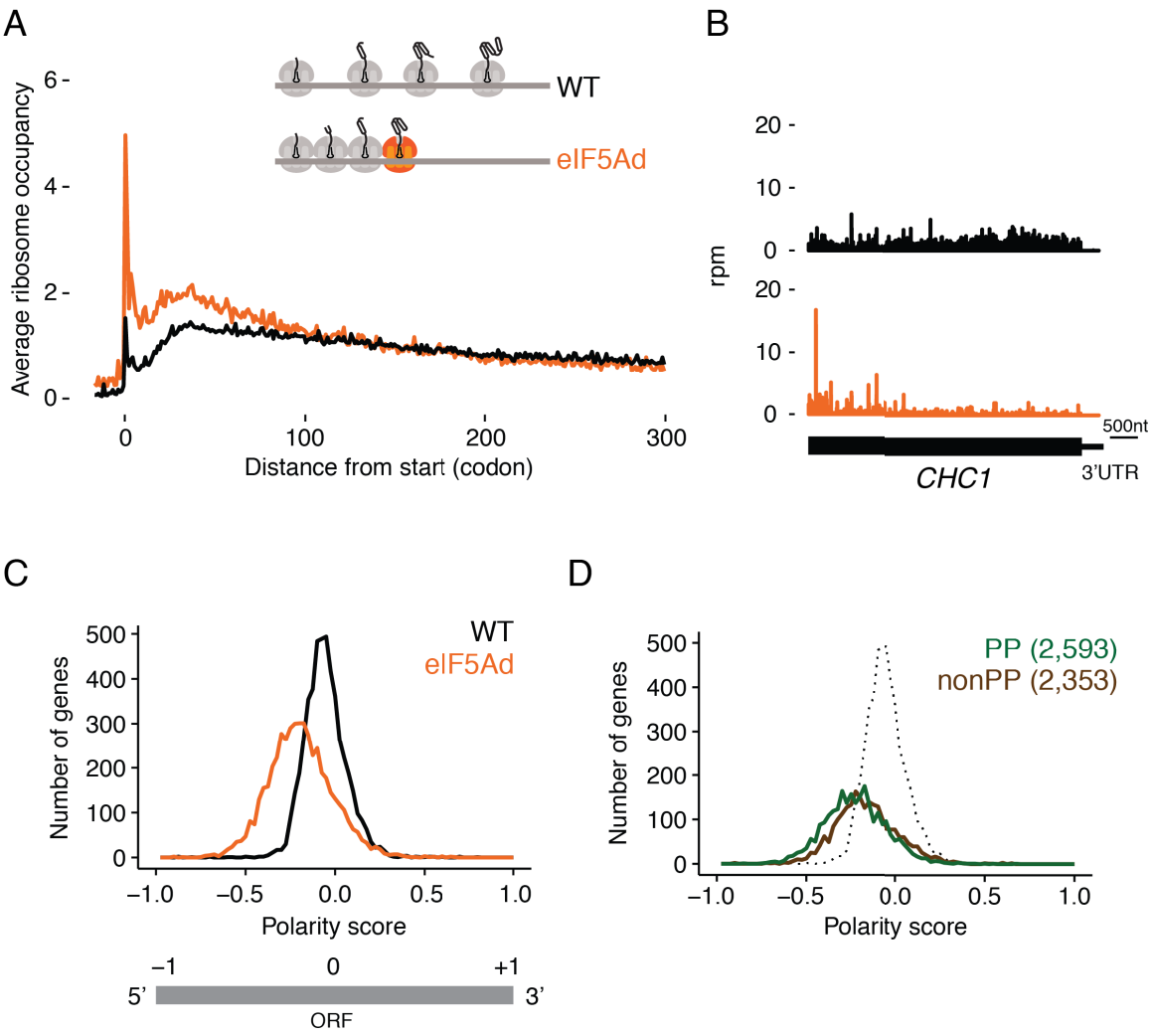
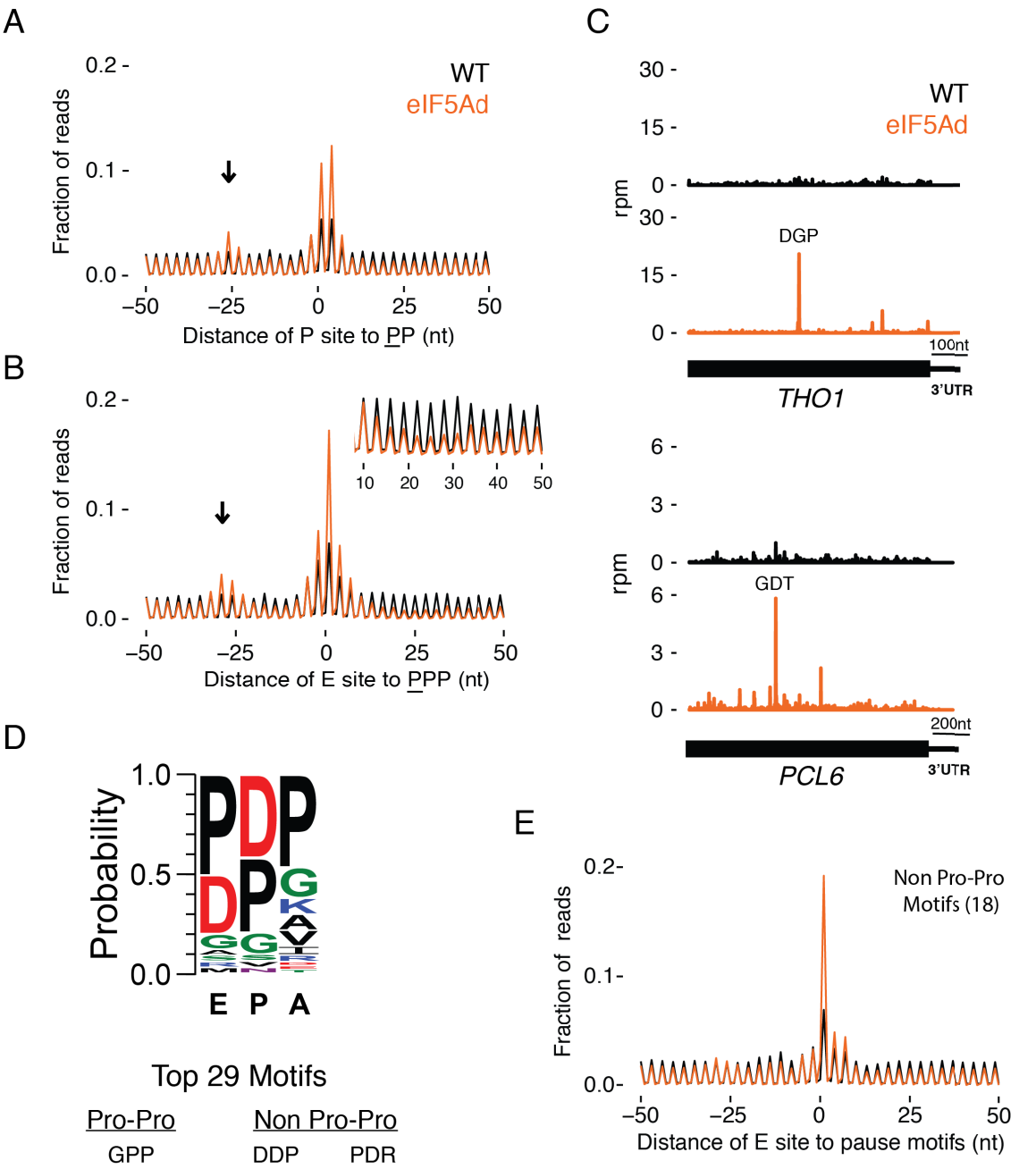

A
Average ribosome occupancy
6
4
2
0
WT
eIF5Ad
0
100
200
300
Distance from start (codon)
B
rpm
20
10
0
20
10
0
500nt
CHC1
3'UTR
C
Number of genes
500
400
300
200
100
0
WT
eIF5Ad
−1.0
−0.5
0.0
0.5
1.0
Polarity score
−1
0
+1
5'
3'
ORF
D
Number of genes
PP (2,593)
nonPP (2,353)
Polarity score

Figure 1



Top 29 Motifs

Pro-Pro	Non Pro-Pro	
GPP	DDP	PDR
DPP	PDP	DPG
PPP	DNP	RDK
SPP	DGP	PDA
PPG	PDK	PDV
PPA	PDG	DVG
PPV	GDP	DDG
PPE	PDI	GGT
PPD	PGP	DSP
APP		
MPP		

Figure 2

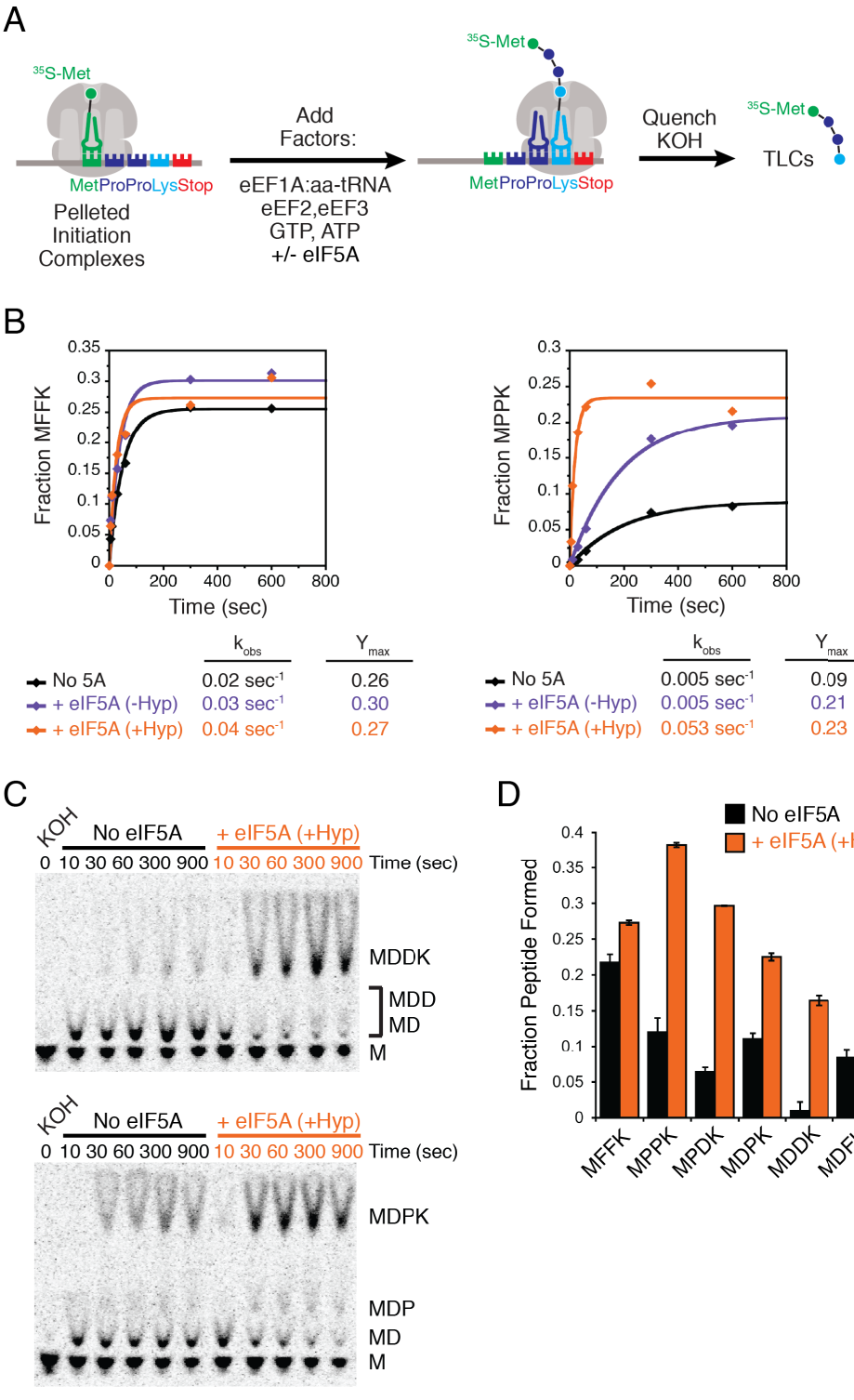

A
35S-Met
MetProProLysStop
Pelleted Initiation Complexes
Add Factors:
eEF1A:aa-tRNA
eEF2,eEF3
GTP, ATP
+/- eIF5A
35S-Met
MetProProLysStop
Quench KOH
35S-Met
TLCs
B
Fraction MFFK
Time (sec)
Fraction MPPK
Time (sec)
kobs
Ymax
No 5A 0.02 sec-1 0.26
+ eIF5A (-Hyp) 0.03 sec-1 0.30
+ eIF5A (+Hyp) 0.04 sec-1 0.27
No 5A 0.005 sec-1 0.09
+ eIF5A (-Hyp) 0.005 sec-1 0.21
+ eIF5A (+Hyp) 0.053 sec-1 0.23
C
KOH
No eIF5A
+ eIF5A (+Hyp)
0 10 30 60 300 900 10 30 60 300 900 Time (sec)
MDDK
MDD
MD
M
KOH
No eIF5A
+ eIF5A (+Hyp)
0 10 30 60 300 900 10 30 60 300 900 Time (sec)
MDPK
MDP
MD
M
D
No eIF5A
+ eIF5A (+Hyp)
Fraction Peptide Formed
MFFK
MPPK
MPDK
MDPK
MDDK
MDFK
MCFK
MCCK

Figure 3

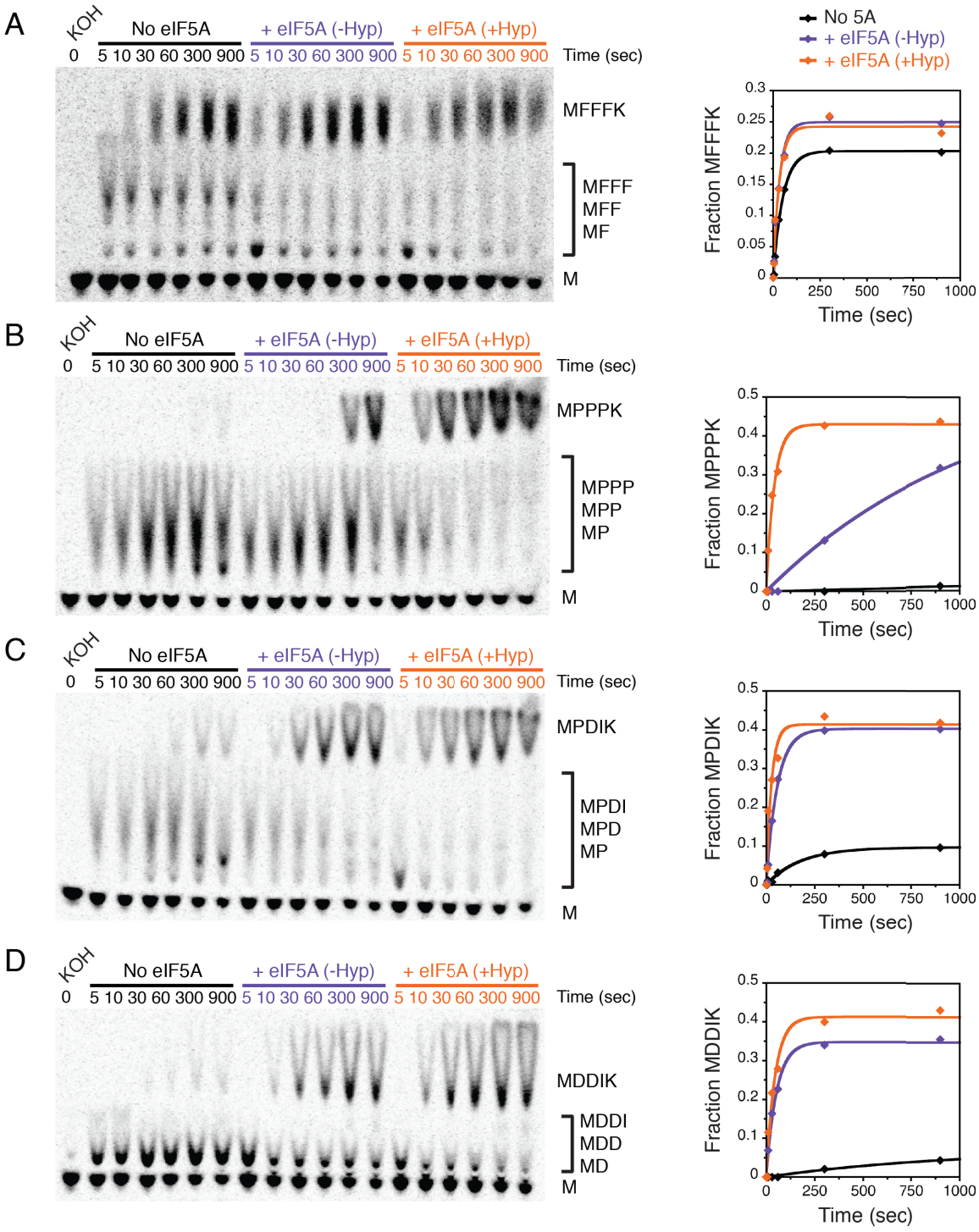

A
KOH
No eIF5A
+ eIF5A (-Hyp)
+ eIF5A (+Hyp)
0 5 10 30 60 300 900 5 10 30 60 300 900 5 10 30 60 300 900
Time (sec)
MFFFK
MFFF
MFF
MF
M
No 5A
+ eIF5A (-Hyp)
+ eIF5A (+Hyp)
Fraction MFFFK
Time (sec)
B
MPPPK
MPPP
MPP
MP
Fraction MPPPK
C
MPDIK
MPDI
MPD
MP
Fraction MPDIK
D
MDDIK
MDDI
MDD
MD
Fraction MDDIK

Figure 4

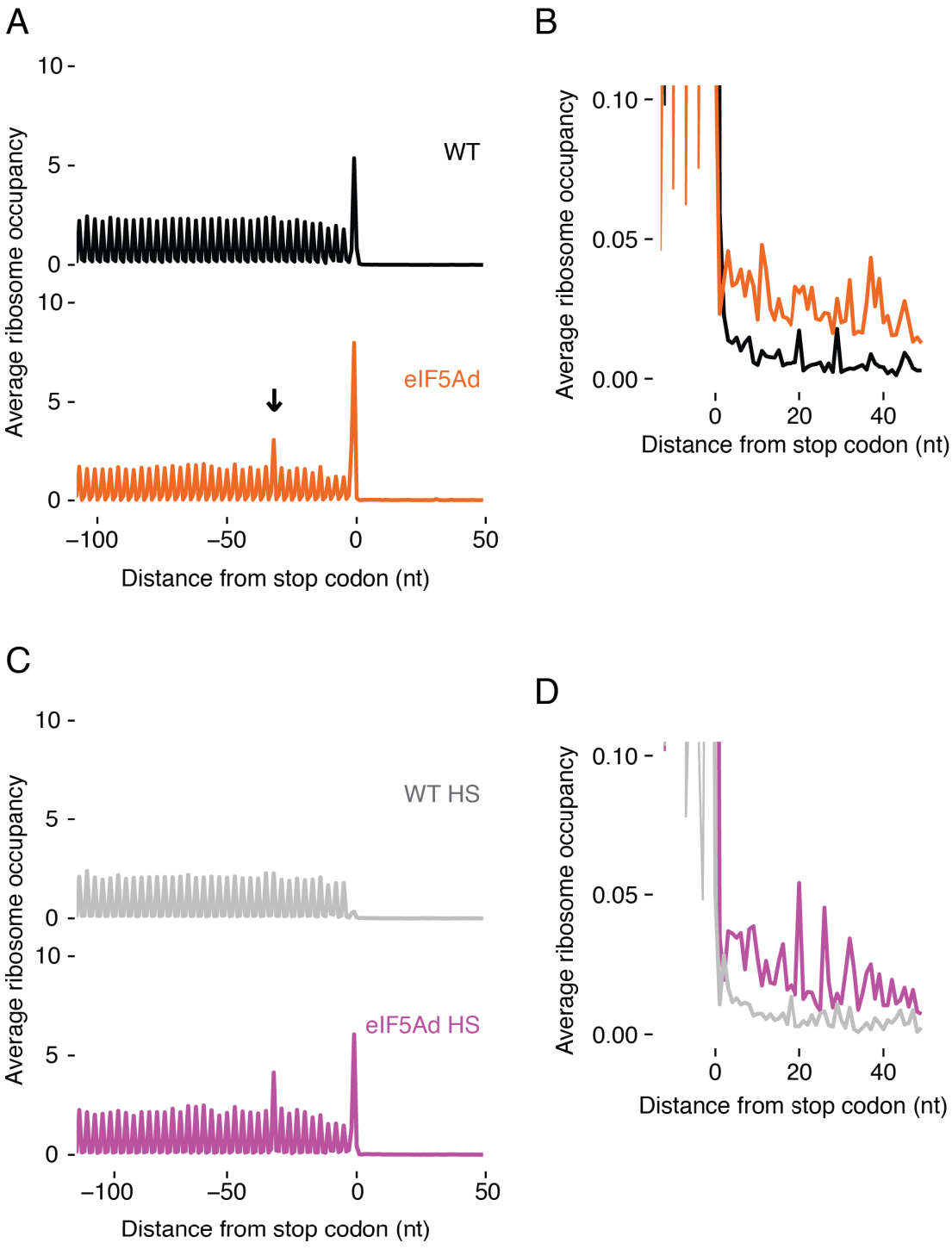

A
B
C
D
WT
eIF5Ad
WT HS
eIF5Ad HS
Average ribosome occupancy
Distance from stop codon (nt)
10
5
0
−100
−50
50
0.10
0.05
0.00
20
40

Figure 5

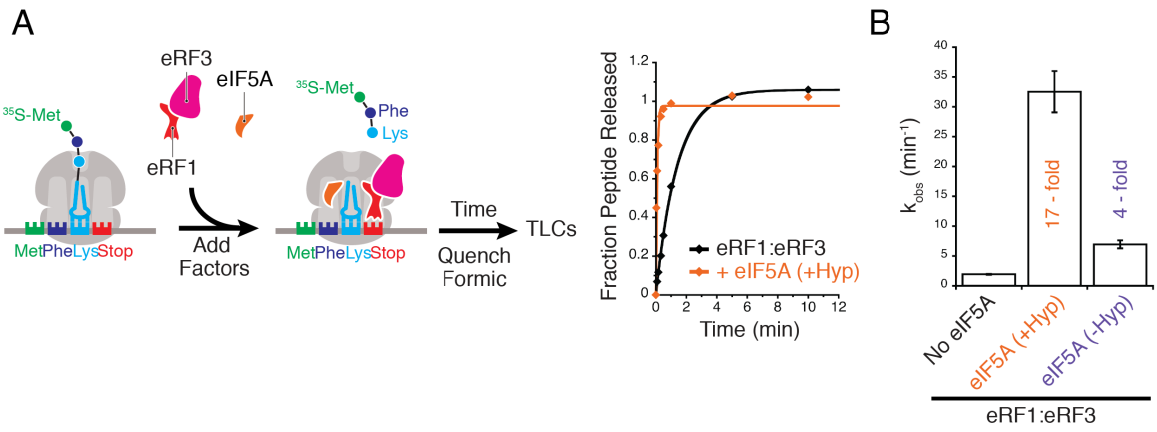

A
eRF3
eIF5A
35S-Met
eRF1
MetPheLysStop
Add
Factors
35S-Met
Phe
Lys
MetPheLysStop
Time
Quench
Formic
TLCs
Fraction Peptide Released
Time (min)
eRF1:eRF3
+ eIF5A (+Hyp)
B
kobs (min-1)
17 - fold
4 - fold
No eIF5A
eIF5A (+Hyp)
eIF5A (-Hyp)
eRF1:eRF3

Figure 6

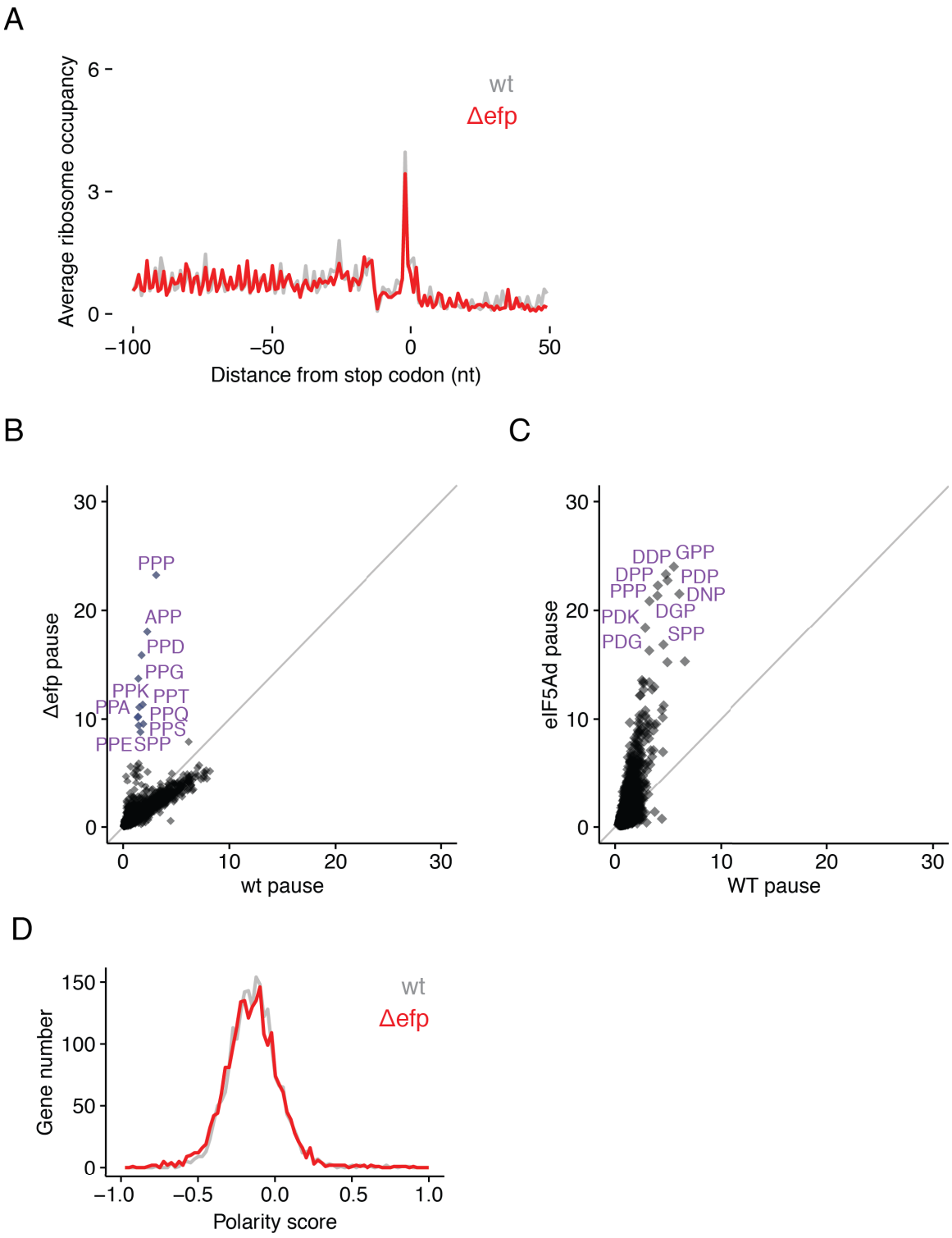
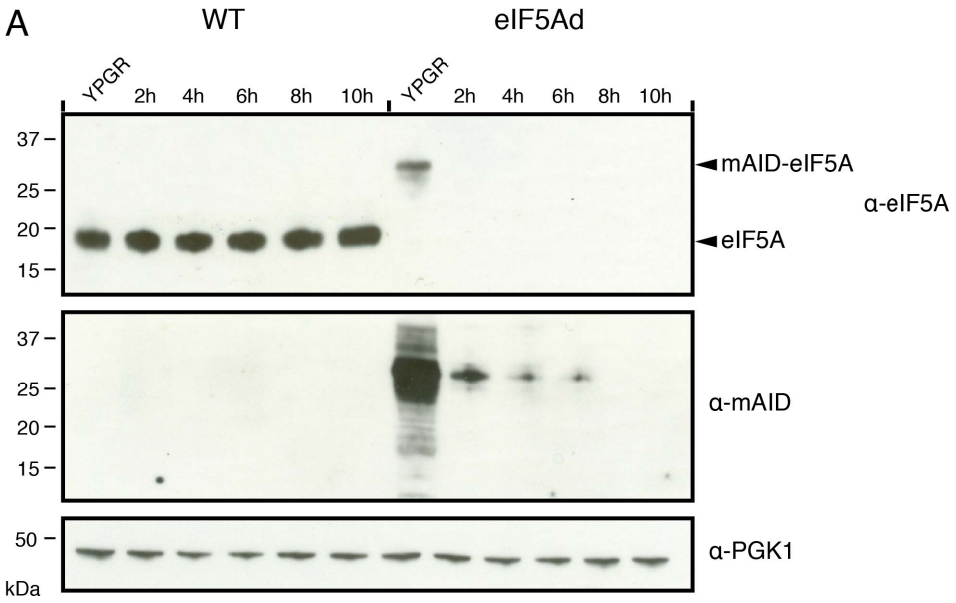
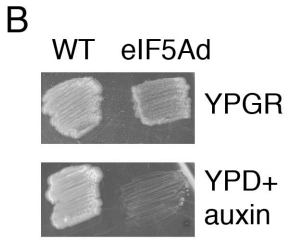
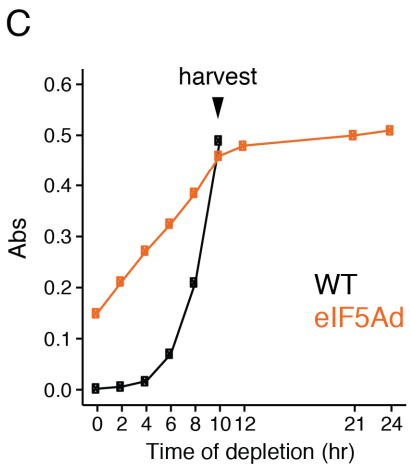
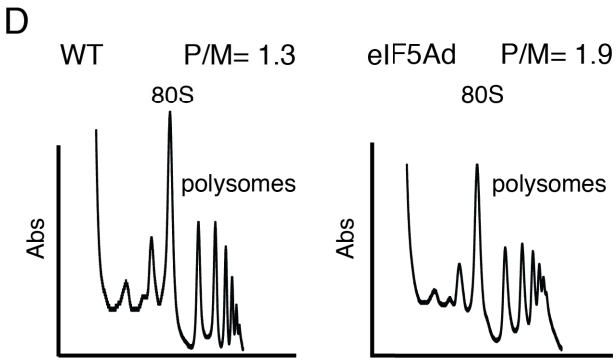

A
Average ribosome occupancy
6
3
0
wt
Δefp
−100
−50
0
50
Distance from stop codon (nt)
B
Δefp pause
30
20
10
0
PPP
APP
PPD
PPG
PPK
PPT
PPA
PPQ
PPS
PPE
SPP
0
10
20
30
wt pause
C
eIF5Ad pause
DDP
GPP
DPP
PDP
PPP
DNP
PDK
DGP
PDG
SPP
WT pause
D
Gene number
150
100
50
0
wt
Δefp
−1.0
−0.5
0.0
0.5
1.0
Polarity score

Figure 7


A
WT
eIF5Ad
YPGR
2h
4h
6h
8h
10h
YPGR
2h
4h
6h
8h
10h
37
25
20
15
mAID-eIF5A
eIF5A
α-eIF5A
37
25
20
15
α-mAID
50
kDa
α-PGK1


B
WT
eIF5Ad
YPGR
YPD+
auxin


C
harvest
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Abs
WT
eIF5Ad
0
2
4
6
8
10
12
21
24
Time of depletion (hr)


D
WT
P/M= 1.3
80S
polysomes
Abs
eIF5Ad
P/M= 1.9
80S
polysomes
Abs

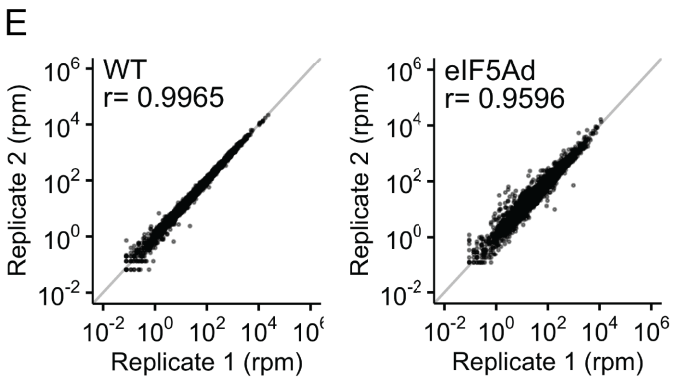

E
WT
r= 0.9965
Replicate 2 (rpm)
Replicate 1 (rpm)
eIF5Ad
r= 0.9596
Replicate 2 (rpm)
Replicate 1 (rpm)

Figure S1

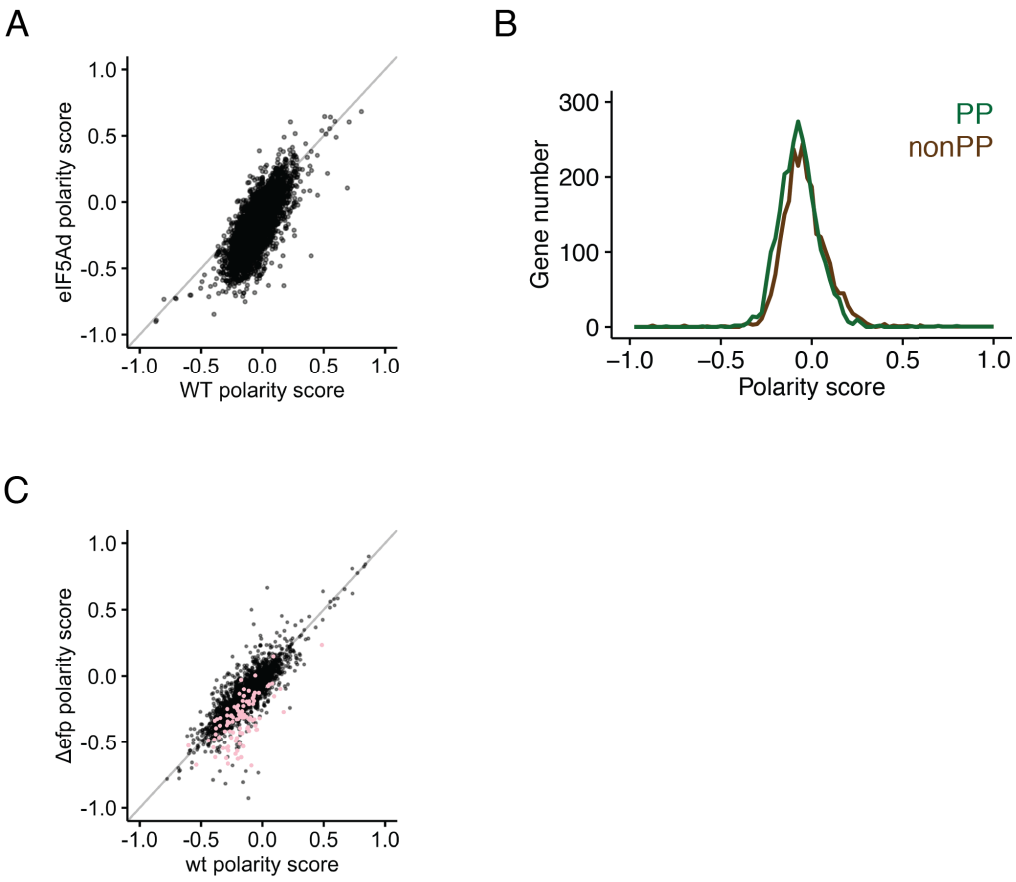

A
eIF5Ad polarity score
WT polarity score
1.0
0.5
0.0
-0.5
-1.0
B
Gene number
Polarity score
PP
nonPP
300
200
100
0
C
Δefp polarity score
wt polarity score

Figure S2

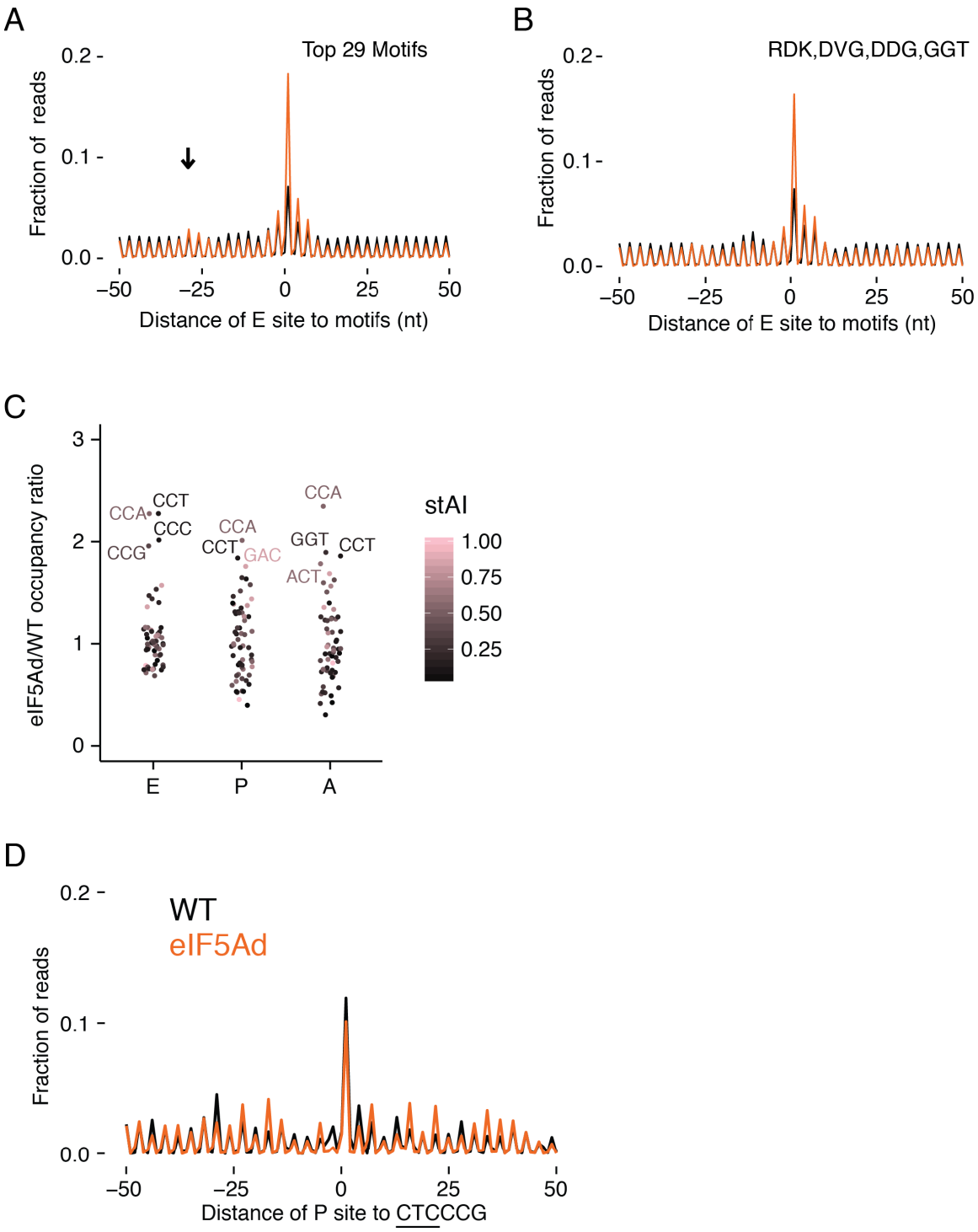

A
Top 29 Motifs
Fraction of reads
0.2
0.1
0.0
−50
−25
0
25
50
Distance of E site to motifs (nt)
B
RDK,DVG,DDG,GGT
Fraction of reads
0.2
0.1
0.0
−50
−25
0
25
50
Distance of E site to motifs (nt)
C
eIF5Ad/WT occupancy ratio
3
2
1
0
CCA
CCT
CCC
CCG
CCA
CCT
GAC
CCA
GGT
CCT
ACT
stAI
1.00
0.75
0.50
0.25
E
P
A
D
Fraction of reads
0.2
0.1
0.0
WT
eIF5Ad
−50
−25
0
25
50
Distance of P site to CTCCCG

Figure S3

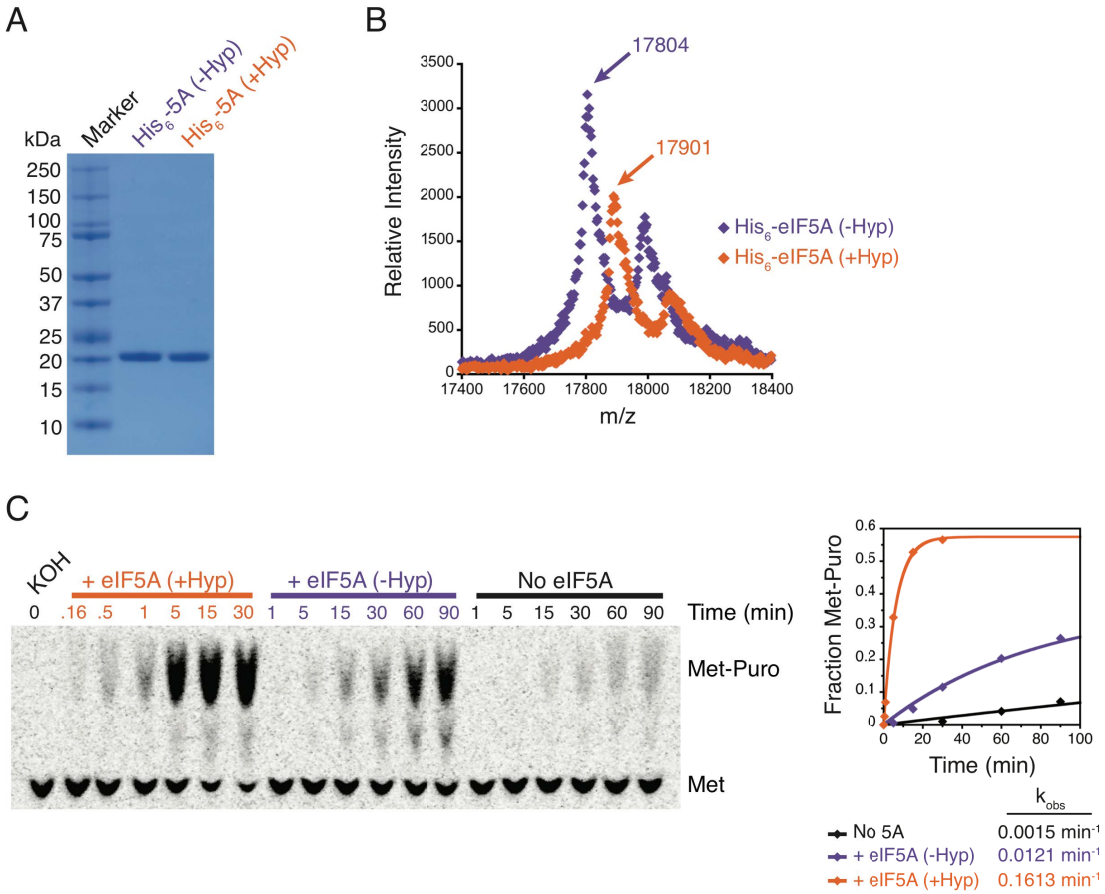

A
kDa
Marker
His6-5A (-Hyp)
His6-5A (+Hyp)
250
150
100
75
50
37
25
20
15
10
B
17804
17901
Relative Intensity
m/z
His6-eIF5A (-Hyp)
His6-eIF5A (+Hyp)
C
KOH
+ eIF5A (+Hyp)
+ eIF5A (-Hyp)
No eIF5A
Time (min)
Met-Puro
Met
Fraction Met-Puro
Time (min)
kobs
No 5A 0.0015 min-1
+ eIF5A (-Hyp) 0.0121 min-1
+ eIF5A (+Hyp) 0.1613 min-1
D
KOH
Formic
No eIF5A
KOH
+ eIF5A (+Hyp)
KOH
Time (sec)
Dropoff products
MDDK
MDD
MD
M
Origin

Figure S4

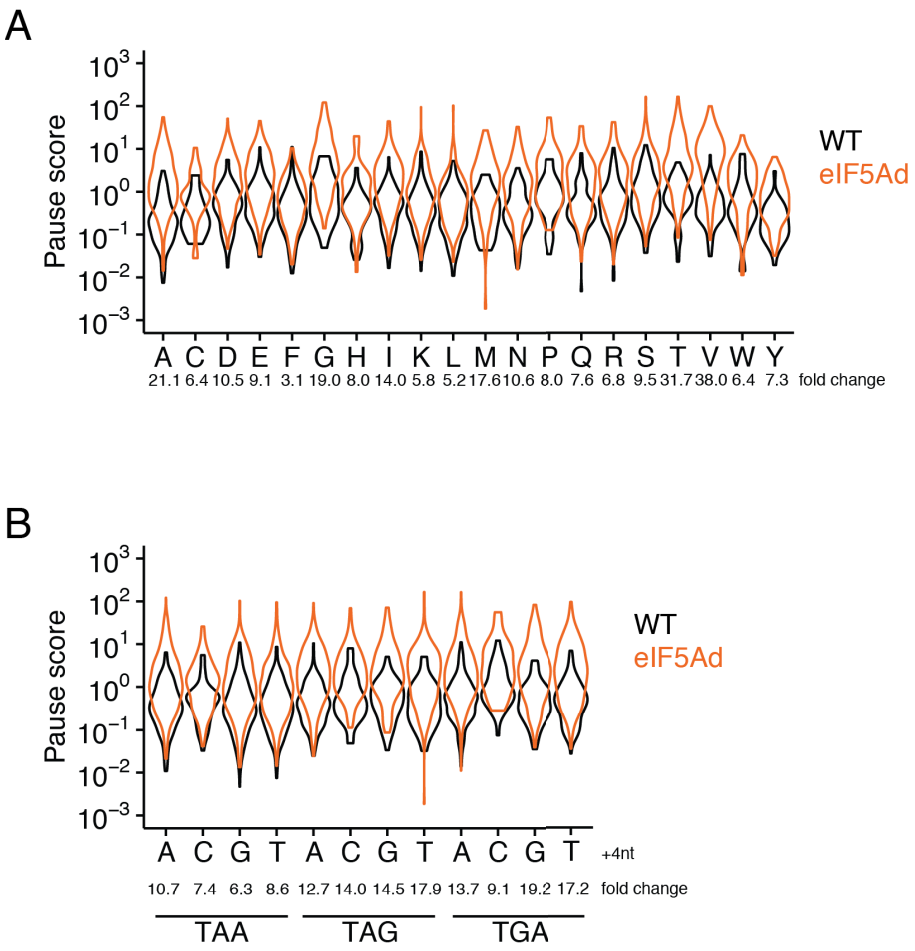
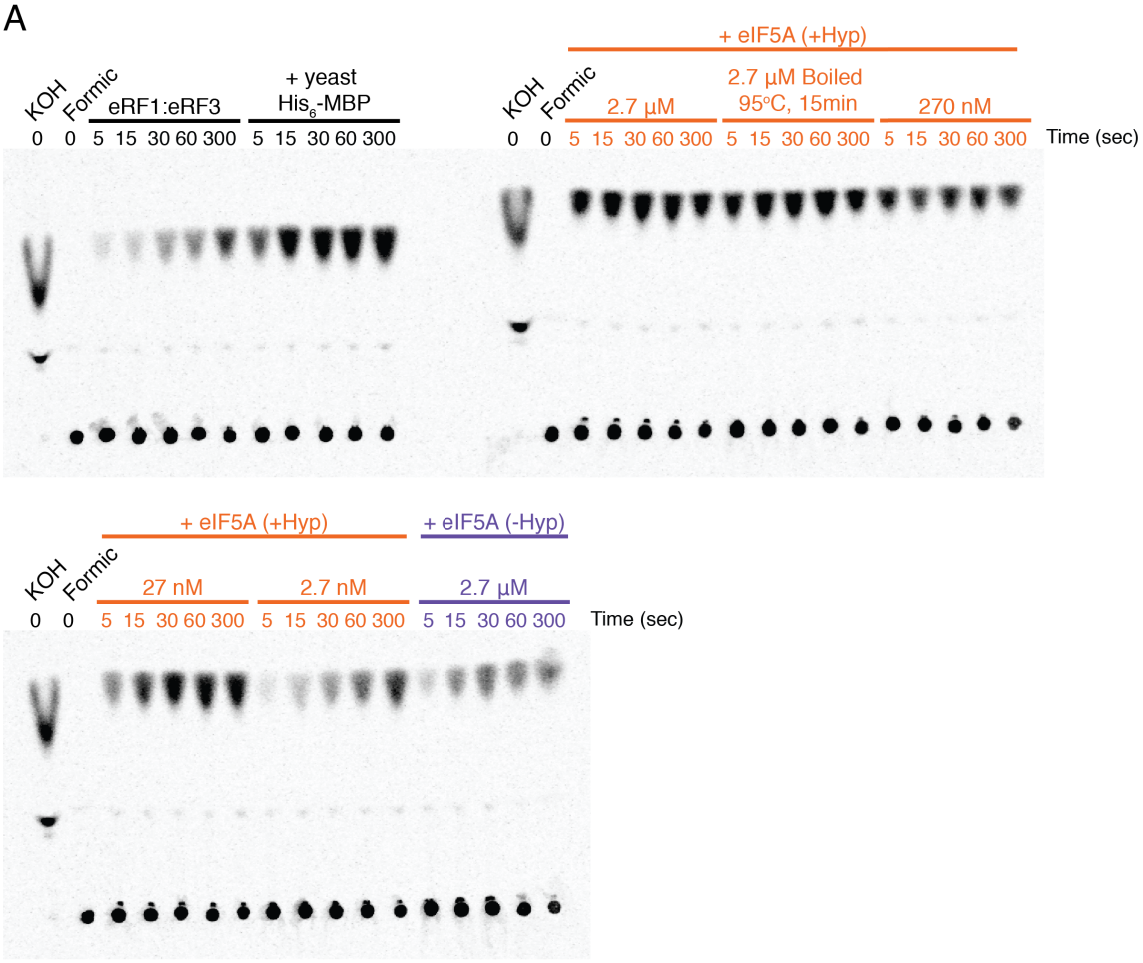

A
Pause score
WT
eIF5Ad
A C D E F G H I K L M N P Q R S T V W Y
21.1 6.4 10.5 9.1 3.1 19.0 8.0 14.0 5.8 5.2 17.6 10.6 8.0 7.6 6.8 9.5 31.7 38.0 6.4 7.3 fold change
B
Pause score
WT
eIF5Ad
A C G T A C G T A C G T +4nt
10.7 7.4 6.3 8.6 12.7 14.0 14.5 17.9 13.7 9.1 19.2 17.2 fold change
TAA
TAG
TGA

Figure S5

A



B

Sample	k_{obs} release (min^{-1})	
eRF1:eRF3	0.5	
+ yeast His_6-MBP	7.8	
+ 2.7 µM eIF5A (+Hyp)	17	Saturating eIF5A
+ 270 nM eIF5A (+Hyp)	15	
+ 27 nM eIF5A (+Hyp)	3.8	
+ 2.7 nM eIF5A (+Hyp)	0.7	
+ 2.7 µM eIF5A (-Hyp)	1.6	
+ 2.7 µM eIF5A (+Hyp) Boiled	13	

Figure S6

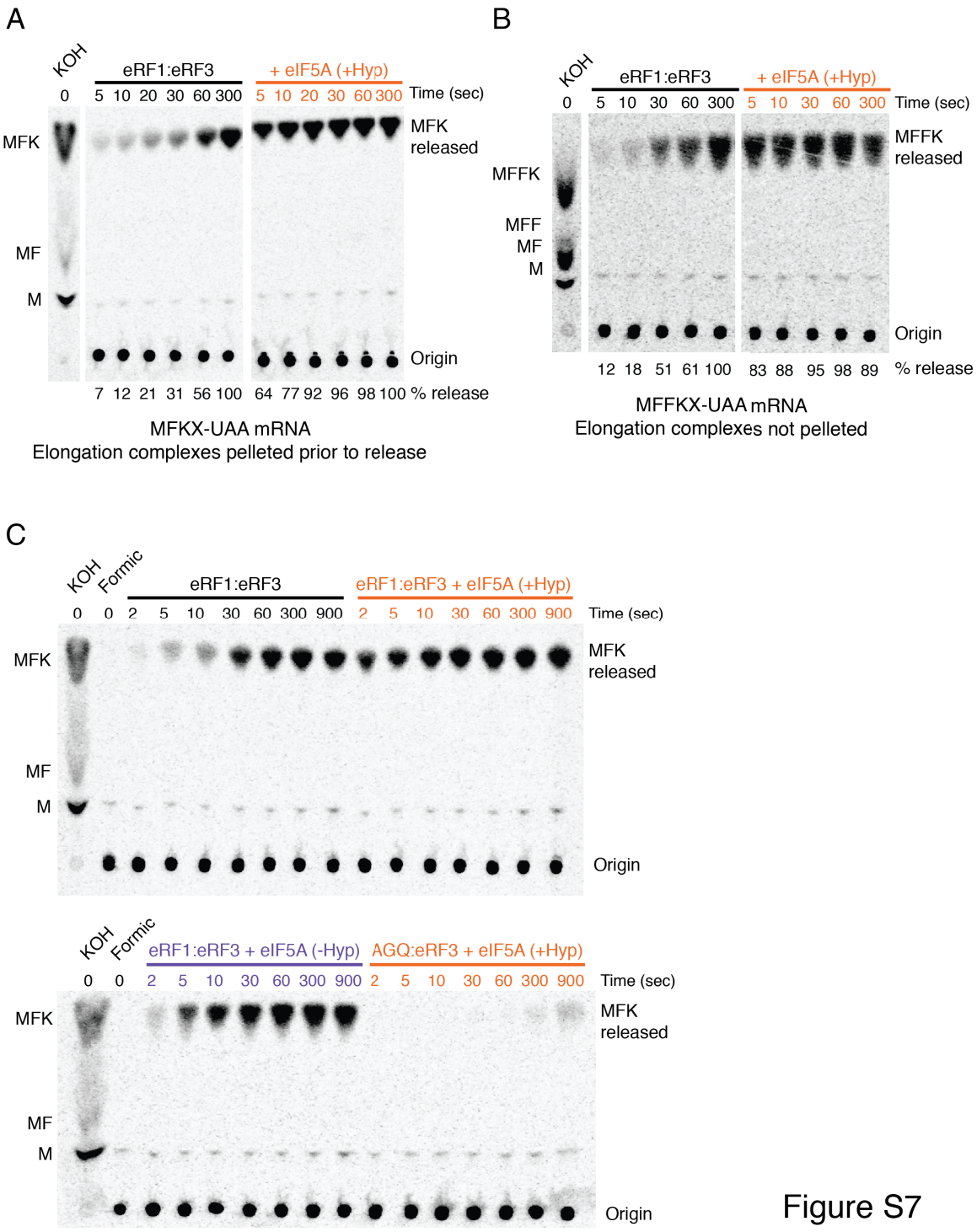

A
KOH
eRF1:eRF3
+ eIF5A (+Hyp)
0
5 10 20 30 60 300
5 10 20 30 60 300
Time (sec)
MFK
MFK released
MF
M
Origin
7 12 21 31 56 100
64 77 92 96 98 100 % release
MFKX-UAA mRNA
Elongation complexes pelleted prior to release
B
KOH
eRF1:eRF3
+ eIF5A (+Hyp)
0
5 10 30 60 300
5 10 30 60 300
Time (sec)
MFFK
MFFK released
MFF
MF
M
Origin
12 18 51 61 100
83 88 95 98 89
% release
MFFKX-UAA mRNA
Elongation complexes not pelleted
C
KOH
Formic
eRF1:eRF3
eRF1:eRF3 + eIF5A (+Hyp)
0 0 2 5 10 30 60 300 900
2 5 10 30 60 300 900
Time (sec)
MFK
MFK released
MF
M
Origin
KOH
Formic
eRF1:eRF3 + eIF5A (-Hyp)
AGQ:eRF3 + eIF5A (+Hyp)
0 0 2 5 10 30 60 300 900
2 5 10 30 60 300 900
Time (sec)
MFK
MFK released
MF
M
Origin

Figure S7

Table S1

Only the top 250 motifs are provided here. The full table is available online with corresponding manuscript Schuller et al. *Molecular Cell* (2017).

Rank	Motif	PauseScore_WT	PauseScore_eIF5Ad	Ratio	Counts_WT	Counts_eIF5Ad
1	GPP	5.531	24.018	4.342	223	227
2	DDP	4.783	23.317	4.875	395	394
3	PDP	4.934	22.751	4.611	212	212
4	DPP	4.028	22.284	5.532	150	153
5	DNP	6.047	21.509	3.557	279	278
6	PPP	3.988	21.349	5.353	832	838
7	DGP	3.228	20.841	6.457	226	227
8	PDK	2.841	18.399	6.476	278	280
9	SPP	4.551	16.852	3.703	434	439
10	PDG	3.221	16.302	5.061	339	343
11	PPG	6.574	15.309	2.329	299	299
12	GDP	4.926	15.234	3.092	186	186
13	PDI	2.531	13.578	5.365	412	415
14	PGP	2.657	13.436	5.058	155	154
15	PDR	3.211	13.389	4.17	150	151
16	PPA	2.66	13.275	4.991	335	339
17	DPG	3.695	12.961	3.508	156	156
18	RDK	3.488	12.73	3.65	339	340
19	PPV	2.67	12.711	4.76	371	377
20	PDA	2.392	12.195	5.099	299	301
21	PDV	2.334	12.173	5.215	359	362
22	PPE	4.578	11.257	2.459	274	277
23	PPD	4.401	10.835	2.462	216	222
24	APP	2.851	10.796	3.787	397	407
25	DVG	2.592	10.405	4.015	337	333
26	DDG	2.293	10.337	4.509	535	533
27	MPP	2.855	10.318	3.614	138	141
28	GGT	4.485	10.148	2.262	427	429
29	DSP	2.321	10.022	4.318	482	476
30	PGR	3.121	9.93	3.181	226	228
31	PPF	1.386	9.678	6.98	179	183
32	PGI	2.176	9.626	4.424	272	275
33	PPR	2.37	9.564	4.035	269	268
34	GDT	3.878	9.549	2.462	359	366
35	TPP	3.245	9.524	2.935	378	379

36	GDG	2.945	9.498	3.225	430	431
37	PGG	3.69	9.496	2.573	258	262
38	PDD	3.079	9.254	3.006	334	333
39	ILW	2.496	9.083	3.639	142	144
40	EPP	2.148	8.985	4.182	251	252
41	NDP	2.376	8.984	3.782	484	488
42	GPG	3.627	8.956	2.469	241	241
43	PDE	2.21	8.835	3.999	416	419
44	PPN	2.877	8.809	3.061	306	309
45	PDQ	2.293	8.744	3.813	163	160
46	GGP	2.683	8.665	3.229	227	231
47	HPP	2.08	8.614	4.142	102	102
48	DGG	3.098	8.608	2.779	349	352
49	NPP	2.471	8.395	3.397	252	257
50	DIG	1.821	8.272	4.542	412	412
51	TGP	3.187	8.216	2.578	245	248
52	DVP	1.968	8.193	4.164	386	390
53	EIG	2.18	7.976	3.658	400	399
54	RPP	2.741	7.952	2.901	248	246
55	DYP	1.852	7.854	4.241	182	182
56	DLP	1.887	7.649	4.053	657	663
57	PPI	1.909	7.642	4.002	334	337
58	EVG	2.114	7.605	3.597	361	357
59	PEP	2.252	7.553	3.355	278	278
60	VPI	1.385	7.537	5.44	471	478
61	PGA	2.26	7.311	3.234	272	273
62	SGP	2.361	7.22	3.058	321	325
63	PPK	2.007	7.218	3.596	394	388
64	SPG	3.617	7.141	1.974	356	362
65	RGP	3.359	7.126	2.121	194	197
66	MDV	1.498	7.039	4.7	192	192
67	GGG	4.507	6.973	1.547	628	640
68	EPG	2.442	6.938	2.841	219	223
69	PPY	2.119	6.897	3.255	172	172
70	PIP	1.505	6.786	4.508	386	388
71	SPT	2.4	6.701	2.793	737	737
72	PGV	1.791	6.693	3.737	314	317
73	RDR	2.531	6.661	2.632	302	303
74	PNI	1.514	6.587	4.349	446	450
75	DAP	1.645	6.568	3.994	286	286
76	SGG	3.485	6.54	1.877	702	710

77	VPF	1.212	6.488	5.351	317	316
78	PNP	1.754	6.33	3.609	281	280
79	GIG	2.036	6.307	3.098	441	447
80	PNE	1.444	6.258	4.333	412	408
81	ADP	2.053	6.231	3.035	250	254
82	DPV	1.741	6.198	3.56	356	358
83	RPR	2.212	6.183	2.795	224	227
84	VPY	1.402	6.113	4.362	225	229
85	EGP	2.313	6.102	2.638	221	224
86	PGD	2.471	6.099	2.468	246	248
87	WLT	1.443	6.095	4.222	121	123
88	EGG	2.192	6.089	2.778	341	341
89	DIP	1.463	6.065	4.147	575	582
90	EIW	3.133	6.063	1.935	118	117
91	GNP	1.876	6.056	3.229	298	296
92	PGL	1.4	6.027	4.306	456	458
93	EDG	1.899	5.965	3.141	637	639
94	SDG	2.565	5.958	2.323	577	579
95	PNV	1.421	5.957	4.193	410	413
96	KPP	2.306	5.933	2.573	289	292
97	EDP	2.086	5.906	2.832	432	438
98	NNP	2.876	5.858	2.036	429	434
99	PVP	1.749	5.816	3.326	408	392
100	GPT	3.31	5.81	1.756	233	238
101	AGP	2.064	5.754	2.788	217	219
102	PGK	2.192	5.737	2.617	287	287
103	VPP	1.972	5.71	2.895	393	396
104	SDP	2.265	5.685	2.51	484	485
105	DPT	1.695	5.678	3.35	465	463
106	ADG	1.947	5.672	2.912	361	360
107	PGN	1.254	5.654	4.507	244	247
108	DSG	1.639	5.629	3.433	530	533
109	PGE	1.967	5.583	2.839	267	272
110	PDN	1.506	5.52	3.664	281	282
111	PPQ	1.866	5.495	2.944	330	332
112	DLG	1.77	5.478	3.096	569	575
113	YGP	2.719	5.475	2.013	207	205
114	GDQ	2.089	5.406	2.588	163	166
115	PAI	1.541	5.394	3.5	367	366
116	DPI	1.172	5.372	4.583	415	416
117	PGQ	1.445	5.307	3.673	187	186

118	TDP	3.01	5.287	1.757	399	402
119	GAP	1.965	5.177	2.634	281	284
120	LPP	2.47	5.151	2.085	860	865
121	VPV	1.544	5.148	3.335	480	486
122	GVG	1.805	5.142	2.849	454	448
123	DYG	1.47	5.121	3.483	205	206
124	PLP	1.778	5.094	2.866	709	714
125	PPM	1.873	5.085	2.715	105	105
126	GDV	1.293	5.059	3.912	493	503
127	DPD	2.092	5.052	2.415	344	351
128	DFP	1.636	5.032	3.075	248	251
129	SPQ	2.065	5.004	2.424	406	405
130	WDT	1.66	4.997	3.011	111	111
131	YDP	1.669	4.992	2.991	309	305
132	DTG	1.467	4.977	3.393	303	309
133	GSP	1.677	4.963	2.96	444	442
134	YIG	1.583	4.949	3.127	244	245
135	DPQ	1.45	4.943	3.409	235	234
136	GVP	1.964	4.856	2.472	406	413
137	NIG	1.351	4.822	3.569	425	426
138	WLP	1.842	4.799	2.606	118	120
139	IPI	1.058	4.78	4.518	573	578
140	YGG	2.614	4.778	1.828	294	296
141	GDI	1.201	4.762	3.964	567	569
142	VPD	2.373	4.76	2.006	356	355
143	GSG	1.557	4.682	3.008	821	825
144	PVG	1.484	4.676	3.15	286	287
145	NVG	1.554	4.618	2.971	426	423
146	PNG	1.342	4.598	3.426	399	406
147	VDW	2.069	4.589	2.218	100	101
148	AVG	1.374	4.582	3.334	460	464
149	YPP	1.8	4.56	2.533	166	167
150	GLP	1.706	4.518	2.648	595	588
151	PEG	1.351	4.497	3.329	353	352
152	QPP	1.945	4.439	2.282	261	263
153	KDK	1.313	4.411	3.36	776	787
154	PNA	1.393	4.391	3.152	345	340
155	AGG	2.215	4.389	1.982	556	552
156	YVG	2.011	4.382	2.178	255	254
157	NGP	2.052	4.351	2.12	274	272
158	PPL	1.164	4.323	3.714	565	574

159	AIG	1.336	4.296	3.216	500	498
160	IPP	1.983	4.28	2.158	472	474
161	DGT	1.915	4.279	2.234	536	538
162	VLQ	1.146	4.27	3.724	605	611
163	TGQ	2.19	4.244	1.938	238	242
164	PDS	1.696	4.241	2.5	487	491
165	DPA	1.836	4.239	2.308	312	312
166	TGG	2.823	4.239	1.501	473	477
167	SDT	2.136	4.231	1.981	586	596
168	SVG	1.388	4.225	3.043	569	567
169	PGT	2.115	4.219	1.995	324	322
170	RGR	2.155	4.21	1.954	306	305
171	VLH	1.109	4.19	3.778	289	287
172	PIG	1.327	4.187	3.156	303	302
173	PGF	1.807	4.184	2.315	262	261
174	GPD	1.749	4.183	2.392	205	206
175	RPK	1.751	4.164	2.378	274	273
176	VVT	1.466	4.156	2.834	510	505
177	ELG	1.881	4.15	2.207	647	645
178	DTP	1.503	4.145	2.758	351	355
179	GIP	1.709	4.138	2.421	508	515
180	GYP	1.641	4.133	2.518	184	185
181	NDG	2.007	4.13	2.058	457	458
182	YDG	2.031	4.129	2.033	225	230
183	VVH	1.401	4.126	2.944	153	154
184	PYP	1.618	4.109	2.539	139	139
185	TLW	1.542	4.103	2.662	123	122
186	VPL	1.135	4.101	3.612	727	732
187	YGE	1.881	4.1	2.18	275	278
188	PDL	1.072	4.099	3.823	602	603
189	PSG	1.486	4.09	2.753	448	453
190	TDT	2.726	4.084	1.498	384	390
191	DRP	2.735	4.077	1.491	179	178
192	ILQ	1.085	4.054	3.735	830	835
193	PAP	1.615	4.047	2.505	385	393
194	TPQ	1.474	4.045	2.744	342	343
195	KDR	1.825	4.032	2.209	408	407
196	DNG	1.581	4.032	2.551	486	489
197	GDA	1.271	4.022	3.165	361	366
198	RDP	1.827	4.021	2.2	259	261
199	TDQ	2.941	4.002	1.361	228	224

200	GEP	1.524	3.995	2.621	266	271
201	PAG	1.556	3.995	2.567	290	291
202	PPH	1.291	3.984	3.087	131	132
203	PDH	1.833	3.983	2.172	112	112
204	ELP	1.754	3.973	2.265	704	712
205	VDK	1.915	3.935	2.054	550	550
206	AVP	1.953	3.931	2.013	341	341
207	ETG	1.591	3.926	2.468	427	426
208	SDQ	2.41	3.92	1.627	283	282
209	MAP	1.593	3.905	2.451	150	151
210	PPT	1.994	3.904	1.958	383	392
211	PDT	1.992	3.893	1.954	286	288
212	TPT	1.478	3.883	2.627	625	611
213	VIH	1.119	3.878	3.467	194	202
214	EIP	1.442	3.873	2.687	500	506
215	GDR	1.7	3.855	2.267	254	251
216	DGD	1.414	3.85	2.723	442	445
217	DEP	1.982	3.818	1.926	366	370
218	ENG	1.679	3.812	2.27	705	700
219	LGP	2.121	3.802	1.793	437	439
220	VIQ	1.12	3.8	3.392	313	312
221	MDP	2.648	3.788	1.431	163	168
222	GNG	1.549	3.777	2.439	498	501
223	VPH	1.218	3.757	3.086	157	162
224	ELW	1.722	3.742	2.173	160	161
225	ILH	0.945	3.738	3.956	345	347
226	VPN	1.133	3.711	3.274	374	376
227	PVI	1.136	3.704	3.261	459	466
228	APG	2.108	3.704	1.757	275	280
229	SPH	2.095	3.69	1.761	181	185
230	GST	1.362	3.69	2.71	757	760
231	DPF	1.317	3.688	2.8	290	292
232	APD	1.801	3.679	2.042	234	235
233	RLK	2.945	3.669	1.246	869	872
234	DHP	1.42	3.656	2.575	147	146
235	ALW	1.879	3.652	1.943	124	122
236	DST	1.183	3.646	3.082	745	746
237	SGT	1.921	3.635	1.892	695	702
238	ADT	1.625	3.631	2.235	348	351
239	AGT	1.799	3.628	2.017	424	422
240	PNN	0.923	3.619	3.92	376	377

241	SIG	1.292	3.618	2.801	658	665
242	MNV	0.954	3.614	3.789	165	166
243	HGG	2.284	3.604	1.578	180	186
244	VVQ	1.467	3.598	2.453	301	304
245	DQP	1.887	3.579	1.897	160	161
246	PYG	1.141	3.564	3.124	183	183
247	HVG	1.367	3.555	2.601	136	135
248	GWI	1.211	3.54	2.922	109	109
249	GAT	1.586	3.529	2.225	476	473
250	PLI	0.886	3.515	3.968	675	676

Table S2

Rank	Motif	PauseScore_WT	PauseScore_eIF5Ad	Ratio	Counts_WT	Counts_eIF5Ad
1	PP	2.871	10.438	3.635	6314	6330
2	DP	2.435	6.192	2.543	6505	6517
3	GP	2.386	5.834	2.445	4354	4369
4	DG	2.03	4.706	2.319	7656	7656
5	PG	2.672	4.454	1.667	4969	4989
6	GG	2.402	4.105	1.709	8309	8325
7	PW	1.75	3.77	2.154	1112	1087
8	VG	1.489	3.625	2.434	7623	7588
9	IG	1.333	3.545	2.658	8261	8258
10	LW	1.531	3.458	2.259	2345	2353
11	VW	1.728	3.203	1.853	1406	1416
12	PD	1.858	3.179	1.711	5701	5713
13	WP	1.62	3.159	1.95	759	766
14	NP	1.952	3.141	1.609	6407	6428
15	LP	1.653	3.024	1.829	12326	12323
16	VP	1.89	3.001	1.588	7383	7360
17	PT	1.562	3.001	1.921	7309	7257
18	IP	1.314	2.967	2.258	8654	8663
19	GT	1.864	2.753	1.477	7900	7854
20	DT	1.445	2.705	1.872	7886	7896
21	GW	1.846	2.679	1.451	1476	1493
22	IW	1.606	2.671	1.663	1903	1904
23	PQ	1.354	2.653	1.96	5009	5032
24	PI	1.081	2.628	2.431	7474	7476
25	DK	1.289	2.55	1.978	9642	9673
26	PV	1.344	2.504	1.863	6871	6840
27	DQ	1.36	2.393	1.76	4561	4584
28	DI	0.956	2.356	2.464	11345	11346
29	LQ	1.105	2.354	2.131	11810	11817
30	DW	1.96	2.334	1.191	1649	1652
31	VT	1.23	2.274	1.849	8610	8530
32	GQ	1.315	2.251	1.712	4357	4360
33	GD	1.264	2.229	1.764	7176	7186
34	DV	1.18	2.175	1.843	9245	9259
35	LG	1.32	2.173	1.647	11794	11776
36	SG	1.344	2.16	1.608	11601	11587
37	PN	1.074	2.158	2.011	6476	6500
38	SP	1.622	2.153	1.327	10252	10266

39	PA	1.428	2.097	1.468	5857	5877
40	YP	1.343	2.071	1.542	3484	3503
41	LT	0.991	2.065	2.085	13938	13942
42	DD	1.091	2.046	1.876	12785	12798
43	AP	1.441	2.026	1.406	6093	6061
44	PR	1.729	2.013	1.164	4597	4609
45	MP	1.604	2.001	1.247	2076	2088
46	MV	1.313	1.962	1.495	3143	3148
47	PE	1.434	1.957	1.365	7676	7686
48	NG	1.296	1.937	1.495	9289	9290
49	DR	1.787	1.926	1.078	5463	5481
50	PY	1.339	1.918	1.433	3567	3577
51	IT	0.869	1.911	2.2	9901	9867
52	VQ	1.13	1.867	1.652	5270	5281
53	IQ	0.949	1.811	1.908	6177	6191
54	LH	0.903	1.797	1.989	5403	5425
55	GI	1.158	1.796	1.551	9093	9071
56	GE	1.307	1.794	1.373	7720	7715
57	TG	1.234	1.777	1.439	8219	8191
58	PF	0.987	1.736	1.758	4961	4969
59	DE	1.119	1.727	1.543	13514	13532
60	YG	1.229	1.727	1.405	4843	4847
61	DA	1.258	1.714	1.362	8371	8370
62	EP	1.592	1.706	1.071	5482	5487
63	PH	1.15	1.703	1.481	2521	2534
64	GA	1.488	1.68	1.129	7394	7407
65	WI	0.921	1.656	1.798	1590	1590
66	MG	1.2	1.649	1.374	3054	3064
67	WG	1.056	1.647	1.559	1312	1316
68	AG	1.261	1.62	1.284	7488	7487
69	YW	1.292	1.561	1.208	980	983
70	TW	1.659	1.559	0.94	1495	1494
71	VD	1.015	1.557	1.534	8959	8973
72	PL	0.863	1.554	1.8	11101	11133
73	SW	1.553	1.53	0.985	2016	2013
74	EG	1.127	1.523	1.352	7402	7407
75	AW	1.596	1.515	0.949	1265	1262
76	GR	1.951	1.51	0.774	5571	5567
77	VH	1.049	1.48	1.41	2924	2944
78	NI	0.835	1.456	1.745	10724	10734
79	LV	0.918	1.425	1.552	13524	13509

80	PK	1.141	1.388	1.217	7627	7625
81	IH	0.884	1.385	1.568	3516	3534
82	WD	0.931	1.378	1.48	1670	1670
83	FP	1.421	1.369	0.964	4863	4874
84	ST	0.949	1.362	1.435	16151	15956
85	GV	1.297	1.353	1.043	7921	7909
86	RP	1.798	1.332	0.741	4363	4369
87	GN	0.809	1.327	1.639	7814	7829
88	TP	1.296	1.287	0.993	7835	7787
89	LD	0.801	1.283	1.602	14397	14403
90	QG	1.251	1.282	1.025	4560	4564
91	TD	1.001	1.276	1.275	7927	7913
92	PS	1.183	1.273	1.076	11122	11102
93	QW	1.383	1.261	0.912	1061	1074
94	QP	1.659	1.255	0.757	4179	4189
95	GL	0.968	1.254	1.295	11971	11954
96	DS	0.987	1.247	1.263	13052	13064
97	MT	0.935	1.24	1.327	3150	3146
98	DL	0.898	1.221	1.359	15148	15132
99	LI	0.768	1.22	1.589	15438	15420
100	GS	1.136	1.216	1.07	12416	12400
101	NE	0.875	1.213	1.386	11652	11673
102	ID	0.801	1.204	1.504	10786	10802
103	WQ	0.743	1.202	1.618	987	993
104	WV	0.995	1.196	1.202	1405	1411
105	HP	1.555	1.191	0.766	2828	2849
106	TQ	0.948	1.189	1.254	5128	5144
107	RW	2.032	1.186	0.584	1215	1219
108	NV	0.928	1.181	1.273	9018	9025
109	LE	0.874	1.18	1.351	16435	16445
110	SQ	0.888	1.179	1.327	9233	9235
111	PM	1.204	1.151	0.956	2026	2038
112	WA	1.099	1.144	1.041	1309	1307
113	ND	0.853	1.139	1.335	10935	10957
114	NW	1.496	1.131	0.756	1626	1644
115	YT	0.875	1.128	1.29	4558	4551
116	MA	1.05	1.107	1.055	3576	3589
117	FW	1.16	1.096	0.944	1349	1332
118	YQ	0.879	1.075	1.224	3421	3434
119	AT	1.113	1.075	0.966	9528	9477
120	NQ	0.846	1.054	1.246	5103	5121

121	LA	1.021	1.051	1.029	13876	13850
122	MS	0.705	1.047	1.485	4989	5013
123	WE	0.876	1.04	1.187	1662	1645
124	WT	0.951	1.037	1.091	1319	1333
125	TT	0.949	1.03	1.085	11214	10938
126	RT	1.32	1.028	0.779	6271	6272
127	FG	1.074	1.028	0.958	6630	6616
128	TE	1.015	1.027	1.012	8592	8482
129	DN	0.695	1.026	1.476	8643	8646
130	EW	1.389	1.016	0.731	1659	1663
131	SD	0.941	1.006	1.069	12254	12272
132	YI	0.693	0.996	1.436	5365	5382
133	HG	1.28	0.986	0.77	2903	2914
134	AD	1.112	0.981	0.882	6925	6931
135	VS	0.978	0.974	0.996	12962	12933
136	YD	0.81	0.973	1.201	5526	5541
137	LN	0.601	0.971	1.615	15408	15419
138	GK	1.053	0.97	0.921	9930	9953
139	QQ	1.004	0.969	0.966	7746	7772
140	VE	1.015	0.969	0.954	9547	9553
141	GH	1.13	0.963	0.853	2897	2906
142	NA	1.047	0.958	0.915	8554	8543
143	VA	1.129	0.953	0.844	8872	8872
144	VN	0.751	0.948	1.262	8191	8226
145	WL	0.693	0.944	1.361	2542	2542
146	LS	0.814	0.942	1.157	22449	22446
147	AI	0.868	0.942	1.085	9727	9718
148	DH	1.106	0.941	0.851	2900	2920
149	NR	1.708	0.939	0.55	5878	5899
150	WK	1.135	0.933	0.822	2305	2311
151	QT	1.035	0.921	0.889	5249	5256
152	MD	0.795	0.915	1.15	3478	3487
153	IA	0.958	0.913	0.952	9505	9515
154	YV	0.928	0.91	0.98	4806	4782
155	MI	0.874	0.909	1.039	3321	3327
156	GF	1.024	0.905	0.883	6117	6106
157	MR	1.434	0.904	0.63	2151	2166
158	AQ	0.962	0.893	0.928	5629	5660
159	WR	1.266	0.893	0.706	1293	1299
160	CE	1.069	0.891	0.834	1600	1610
161	GY	1.157	0.888	0.767	4672	4655

162	IV	0.824	0.887	1.076	9808	9794
163	YR	1.387	0.887	0.64	3704	3709
164	MK	0.984	0.882	0.896	4096	4112
165	YH	0.775	0.872	1.124	2114	2117
166	YA	1.031	0.871	0.845	4418	4431
167	DY	1.134	0.865	0.763	5541	5555
168	FQ	0.885	0.864	0.976	5115	5134
169	RG	1.351	0.861	0.637	5559	5578
170	TI	0.764	0.861	1.126	9925	9877
171	VI	0.876	0.86	0.983	9458	9454
172	TV	1.016	0.86	0.847	9340	9289
173	TA	1.102	0.86	0.78	9241	9195
174	KP	1.463	0.856	0.585	7816	7833
175	CP	1.638	0.85	0.519	1337	1340
176	PC	1.215	0.845	0.696	949	951
177	FT	0.846	0.843	0.997	6651	6633
178	RQ	1.151	0.843	0.732	4570	4582
179	RI	1.189	0.841	0.707	7720	7745
180	WH	0.737	0.836	1.134	529	531
181	TN	0.616	0.831	1.35	9150	9096
182	LY	0.781	0.808	1.035	7984	7992
183	SE	0.928	0.808	0.87	12705	12656
184	QD	0.948	0.801	0.845	5667	5678
185	IE	0.764	0.801	1.049	10972	10971
186	WS	1.058	0.798	0.754	2206	2208
187	LK	0.871	0.796	0.913	20430	20457
188	MQ	0.763	0.794	1.041	1847	1854
189	VV	0.926	0.794	0.858	9253	9241
190	FD	0.765	0.793	1.037	7212	7223
191	FR	1.522	0.79	0.519	4240	4246
192	TH	0.818	0.788	0.964	3010	3042
193	ME	0.884	0.778	0.88	3512	3528
194	MM	1.047	0.777	0.742	1130	1140
195	YE	0.793	0.776	0.979	5358	5355
196	AV	1.04	0.776	0.746	8361	8350
197	IN	0.565	0.773	1.368	10286	10304
198	LR	1.483	0.771	0.52	11872	11890
199	WN	0.669	0.763	1.142	1715	1718
200	SA	1.054	0.762	0.724	12609	12538
201	FH	0.833	0.761	0.914	2709	2710
202	WY	1.088	0.761	0.699	930	936

203	EH	0.847	0.759	0.897	3277	3289
204	FI	0.766	0.759	0.99	8020	8027
205	IS	0.747	0.758	1.014	15848	15809
206	DM	1	0.757	0.756	2879	2891
207	MF	0.685	0.753	1.098	2162	2168
208	MN	0.544	0.749	1.376	3458	3474
209	EQ	0.839	0.748	0.891	6934	6963
210	RV	1.043	0.743	0.713	5843	5854
211	NL	0.75	0.739	0.985	13884	13917
212	NT	0.811	0.733	0.905	9113	9132
213	II	0.65	0.732	1.127	11377	11361
214	AK	1.031	0.726	0.704	10637	10660
215	IR	1.387	0.723	0.522	7409	7413
216	NS	0.868	0.718	0.827	15590	15593
217	VR	1.572	0.715	0.455	6049	6058
218	RD	1.216	0.715	0.588	6472	6470
219	LF	0.665	0.713	1.071	10984	10992
220	ET	0.878	0.708	0.807	9722	9695
221	AA	1.131	0.705	0.623	10711	10694
222	HQ	0.884	0.697	0.788	2150	2167
223	HV	0.921	0.695	0.755	3031	3045
224	AR	1.362	0.694	0.51	6353	6371
225	ED	0.701	0.693	0.988	11865	11873
226	GM	1.175	0.691	0.588	2485	2491
227	AE	1.024	0.69	0.673	8178	8211
228	HW	1.206	0.685	0.568	641	643
229	YL	0.66	0.681	1.031	8670	8657
230	LL	0.701	0.679	0.97	25523	25484
231	TR	1.228	0.677	0.551	6185	6211
232	CG	1.193	0.676	0.567	2115	2118
233	HE	0.934	0.675	0.722	3374	3378
234	NK	0.771	0.67	0.869	11230	11262
235	FA	0.947	0.666	0.703	6311	6305
236	SS	0.869	0.666	0.766	29487	29110
237	KG	1.297	0.664	0.512	7939	7937
238	HT	0.969	0.661	0.682	2778	2788
239	LM	0.823	0.661	0.803	4532	4538
240	EV	0.957	0.648	0.677	9012	8997
241	SI	0.798	0.647	0.81	15054	15055
242	FK	0.933	0.645	0.692	8186	8200
243	GC	1.291	0.645	0.5	1842	1848

244	SR	1.396	0.643	0.46	10195	10203
245	FV	0.766	0.643	0.839	6555	6573
246	VY	0.96	0.642	0.669	4791	4799
247	TL	0.64	0.641	1	14537	14564
248	YY	0.783	0.638	0.815	3597	3606
249	RL	0.843	0.638	0.756	11192	11215
250	RR	1.651	0.637	0.386	7476	7500
251	HD	0.911	0.637	0.699	3260	3274
252	QV	1.093	0.631	0.578	5341	5336
253	QI	0.847	0.63	0.744	6437	6447
254	TC	1.088	0.63	0.579	1644	1649
255	FE	0.773	0.629	0.814	7385	7377
256	RK	1.237	0.624	0.505	9876	9872
257	YK	0.886	0.623	0.704	5869	5890
258	NH	0.833	0.623	0.748	3112	3133
259	RY	1.228	0.618	0.503	4072	4075
260	RM	1.074	0.617	0.575	2106	2125
261	TS	0.851	0.616	0.724	15315	15124
262	WF	0.715	0.616	0.862	1315	1320
263	AY	1.09	0.616	0.565	4440	4442
264	EI	0.681	0.613	0.901	11554	11548
265	DF	0.845	0.61	0.721	7212	7217
266	QA	1.181	0.609	0.516	5182	5202
267	NY	0.983	0.608	0.618	5224	5246
268	CR	1.743	0.608	0.349	1266	1264
269	MC	1.059	0.602	0.569	629	628
270	YS	0.873	0.6	0.688	7177	7178
271	SH	0.791	0.6	0.759	4985	5002
272	KW	1.306	0.599	0.459	1996	2008
273	SV	0.879	0.599	0.681	12311	12271
274	HI	0.766	0.598	0.78	3603	3621
275	KV	0.882	0.595	0.674	10415	10411
276	NN	0.523	0.595	1.137	12884	12902
277	MY	0.808	0.593	0.733	1457	1461
278	SY	0.912	0.593	0.65	6553	6569
279	TY	0.958	0.589	0.615	4358	4361
280	YN	0.587	0.584	0.994	5062	5079
281	IK	0.8	0.583	0.729	12137	12132
282	LC	0.921	0.582	0.632	3299	3299
283	VL	0.843	0.578	0.685	14106	14096
284	QR	1.397	0.575	0.412	5327	5349

285	CD	0.965	0.575	0.596	1719	1731
286	QE	0.995	0.572	0.575	7439	7454
287	CI	0.822	0.571	0.694	2405	2414
288	QS	0.892	0.571	0.64	7661	7695
289	IL	0.62	0.569	0.918	16974	16987
290	SN	0.558	0.569	1.021	15621	15619
291	VM	0.998	0.567	0.568	2691	2694
292	HR	1.501	0.567	0.378	2531	2546
293	RE	1.03	0.567	0.55	7429	7435
294	AS	0.882	0.565	0.641	13245	13187
295	ER	1.232	0.563	0.457	8003	8025
296	FY	0.781	0.563	0.72	4124	4129
297	WC	1.165	0.563	0.483	582	588
298	AN	0.687	0.562	0.817	8294	8310
299	VC	1.386	0.56	0.404	1948	1925
300	AL	0.691	0.558	0.808	14205	14189
301	QH	0.969	0.555	0.573	2195	2201
302	IY	0.769	0.555	0.721	5434	5437
303	HK	1.063	0.55	0.518	3328	3348
304	QL	0.775	0.549	0.708	10798	10827
305	VF	0.757	0.545	0.72	6440	6433
306	VK	0.945	0.542	0.573	10304	10292
307	RA	1.304	0.537	0.412	6002	6021
308	EA	1.004	0.535	0.533	9191	9192
309	AH	0.92	0.535	0.581	2853	2874
310	FL	0.64	0.534	0.835	11143	11136
311	TF	0.746	0.531	0.711	6672	6660
312	RH	1.162	0.526	0.453	2589	2603
313	CL	0.753	0.525	0.698	3337	3343
314	SL	0.718	0.521	0.726	22724	22722
315	AM	1.024	0.518	0.506	3012	3011
316	RS	1.073	0.511	0.476	9645	9667
317	QY	0.984	0.511	0.519	3562	3577
318	QK	0.977	0.509	0.521	7315	7316
319	EE	0.76	0.508	0.669	16302	16323
320	KD	0.86	0.506	0.588	11305	11308
321	ES	0.724	0.504	0.695	12936	12858
322	FS	0.824	0.5	0.607	9493	9505
323	TM	1.03	0.488	0.473	2662	2665
324	QN	0.658	0.481	0.731	6281	6313
325	FM	0.953	0.481	0.505	2079	2073

326	KE	1.001	0.48	0.479	14519	14524
327	YM	0.857	0.479	0.559	1645	1654
328	EY	0.797	0.479	0.6	5841	5834
329	IC	0.903	0.478	0.53	2369	2378
330	IF	0.592	0.478	0.808	7949	7958
331	KQ	0.938	0.471	0.502	7417	7443
332	HA	1.05	0.471	0.448	2592	2612
333	KR	1.291	0.465	0.36	10628	10639
334	IM	0.757	0.464	0.613	3148	3165
335	FN	0.551	0.464	0.842	6942	6963
336	CY	1.116	0.464	0.415	1150	1164
337	ML	0.654	0.461	0.704	4635	4651
338	DC	1.063	0.458	0.431	1689	1683
339	QF	0.773	0.457	0.592	4473	4492
340	NM	0.708	0.455	0.643	2818	2831
341	CQ	0.81	0.453	0.559	1195	1203
342	TK	0.781	0.447	0.572	10792	10787
343	EL	0.712	0.447	0.628	17001	16990
344	RF	1.018	0.447	0.439	5003	5021
345	YC	0.981	0.447	0.455	1376	1381
346	SK	0.803	0.444	0.553	17305	17290
347	RN	0.725	0.443	0.611	7055	7077
348	KI	0.706	0.443	0.627	13125	13149
349	RC	1.524	0.443	0.291	1459	1477
350	CA	1.056	0.443	0.419	1601	1604
351	EK	0.785	0.441	0.562	14548	14563
352	HY	0.818	0.438	0.535	2021	2026
353	HL	0.665	0.432	0.651	5567	5593
354	QM	0.868	0.432	0.498	1965	1981
355	YF	0.617	0.431	0.7	4225	4240
356	KL	0.751	0.43	0.573	19200	19190
357	HH	0.874	0.424	0.485	1763	1780
358	NF	0.787	0.423	0.537	6909	6912
359	QC	0.885	0.422	0.477	1046	1059
360	NC	0.973	0.419	0.431	1815	1830
361	KT	0.82	0.417	0.508	10532	10534
362	FF	0.64	0.414	0.646	5608	5609
363	EN	0.499	0.406	0.813	12210	12211
364	CK	0.898	0.406	0.452	1920	1928
365	SF	0.729	0.402	0.551	10050	10029
366	SC	1.147	0.398	0.347	2503	2501

367	FC	0.912	0.398	0.436	1408	1402
368	KK	1.062	0.398	0.375	18253	18276
369	CV	0.821	0.395	0.481	1885	1881
370	CM	0.906	0.391	0.432	553	554
371	SM	0.834	0.391	0.469	4512	4522
372	MH	0.715	0.388	0.543	954	953
373	CH	0.932	0.388	0.416	839	841
374	KY	0.728	0.386	0.531	7030	7042
375	KA	0.928	0.386	0.416	9733	9734
376	HM	0.891	0.382	0.429	938	945
377	HS	0.892	0.379	0.425	5098	5122
378	EC	0.833	0.375	0.451	1694	1700
379	AF	0.807	0.372	0.461	5968	5959
380	HN	0.59	0.368	0.624	3059	3080
381	CS	0.883	0.36	0.408	2754	2743
382	KH	0.766	0.347	0.453	3846	3874
383	CN	0.648	0.343	0.529	1517	1523
384	AC	1.046	0.342	0.327	1670	1668
385	HC	0.882	0.339	0.384	807	811
386	CT	0.894	0.336	0.376	1617	1601
387	EF	0.655	0.319	0.487	7268	7278
388	KC	1.033	0.316	0.306	2162	2166
389	EM	0.81	0.313	0.387	3352	3361
390	HF	0.643	0.308	0.478	2509	2524
391	KS	0.714	0.298	0.417	14971	14984
392	KF	0.643	0.295	0.459	8764	8778
393	KM	0.674	0.281	0.417	3287	3298
394	CF	0.704	0.268	0.382	1666	1654
395	CC	1	0.267	0.267	689	694
396	KN	0.54	0.251	0.465	12141	12159

Table S3

Name	Identifier	Sequence (5' --> 3')
Oligo for tRNAPro(UGG)	oAS 402	CCAAAGCGAGAATCATACCACTAGAC – 3' Biotin-TEG
Oligo for tRNAAsp(GUC)	oAS 415	GACAAGCGCCCATTCTGACCATTAAAC – 3' Biotin-TEG
Oligo for tRNAIle(AAU)	oAS 429	ATTAGCACGGTGCCTTAACCAACTGGGC – 3' Biotin-TEG
T7 Clamp	oAS 145	TAATACGACTCACTATAGG
MFFFKX-UAA	oAS 445	AGAGAGAGAGAGTTATTTGAAGAAGAAC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MPPPKX-UAA	oAS 405	AGAGAGAGAGAGTTATTTTGGTGGTGGC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MPDIKX-UAA	oAS 440	AGAGAGAGAGAGTTATTTAATGTCTGGC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MDDIKX-UAA	oAS 546	AGAGAGAGAGAGTTATTTAATGTCGTCC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MFFKX-UAA	oAS 408	GAGAGAGAGAGAGAGTTATTTGAAGAAC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MPPKX-UAA	oAS 404	GAGAGAGAGAGAGAGTTATTTTGGTGGC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MPDKX-UAA	oAS 423	GAGAGAGAGAGAGAGTTATTTGTCTGGC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MDPKX-UAA	oAS 422	GAGAGAGAGAGAGAGTTATTTTGGGTCC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MDDKX-UAA	oAS 424	GAGAGAGAGAGAGAGTTATTTGTCGTCC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MDFKX-UAA	oAS 447	GAGAGAGAGAGAGAGTTATTTGAAGTCC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MCFKX-UAA	oAS 450	GAGAGAGAGAGAGAGTTATTTGAAGCAC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MCCKX-UAA	oAS 449	GAGAGAGAGAGAGAGTTATTTGCAGCAC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA
MFKX-UAA	oAS 397	AGAGAGAGAGAGAGAGAGTTATTTGAAC ATAGAGAGAGAGAGAGATTCCTATAGTG AGTCGTATTA

HYP2KO_F	oCW 01	CTTTCTAACGATTCTACTTCTGTAGCCAA TTACTCATAGACTCCCAAACACACACAAA TACCAACTCATATATACAGCATAGGCCAC TAGTGGATCTG
HYP2KO_R	oCW 02	CATGATGTGATGGAAGGGGCGTCGGAGT TTTTTTTTCTTTTTTCATTTATATCCCATGC CATGATGTTAACCGGTAGCTGAAGCTTCG TACG
HO-osTIR1_F	oCW 03	GTATTCAATTCCTATTCTAAATGGCTTTTA TTTCTATTACAACTATTAGCTCTAAATCC ATATCCTCATAAGCAGCAATCAATTCTAT CTATACTTTAAAGGGTGTACAATATGGAC TTCCTC
HO-Ura3p_R	oCW 04	GTATCTTCCATAGCATCTAGCACATACTC GATTTTTACCACTCCAATCTTTATAAAAA TACTTGATTCCCTTTCTGGGACAAGCAAC ACAGTGTTTTAGATTCTTCAATTCATCATT TTTTTTTTATTC

Chapter III: Directed hydroxyl radical probing reveals Upf1 binding to the 80S ribosomal E site rRNA at the L1 stalk

ABSTRACT

Upf1 is an SF1-family RNA helicase that is essential for nonsense-mediated decay (NMD) in eukaryotes. While Upf1 has been shown to interact with 80S ribosomes, the molecular details of this interaction were unknown. Using purified proteins and high-throughput sequencing combined with Fe-BABE directed hydroxyl radical probing (HTS-BABE) we have characterized the interaction between Upf1 and the 80S ribosome. We identify the 1C domain of Upf1, an alpha-helical insertion in the RecA helicase core, to be essential for ribosome binding, and determine that the L1 stalk of 25S rRNA is the binding site for Upf1 on the ribosome. Using the cleavage sites identified by radical probing and high-resolution structures of both Upf1 and the human 80S ribosome, we are able to provide a model of the Upf1:80S interaction. Our model requires that the L1 stalk adopt an open configuration as adopted by an un-rotated, or classical-state, ribosome. Our results shed light on the interaction between Upf1 and the ribosome, and suggest a potential role for Upf1 in stabilization of the ribosome classical state through binding to the L1 stalk.

INTRODUCTION

In eukaryotes there are several mRNA surveillance pathways that selectively degrade mRNAs and rescue ribosomes on mRNAs that are unlikely to produce biologically relevant products (Shoemaker and Green, 2012). One of these pathways, nonsense-mediated decay (NMD), selectively degrades mRNAs that contain what is broadly referred to as a premature termination codon (PTC). These mRNAs are created by multiple routes: genes may carry a mutation that results in a PTC (Holbrook et al., 2004), inefficient splicing may lead to export of a pre-mRNA with a PTC (almost inevitably) encoded in the intron (He et al., 1993; Mitrovich and Anderson, 2000), or the stop codons of upstream open reading frames (Mendell et al., 2004; Welch and Jacobson, 1999) and non-coding RNAs (Marquardt et al., 2011; Smith et al., 2014) can be sensed as PTCs. In these situations, the cell recognizes an mRNA with a PTC as a substrate for degradation, while those mRNAs with normal termination codons (TC) are maintained. This discrimination is still only partially understood, though NMD has been shown *in vivo* to regulate 1-10% of the transcriptome, demonstrating that NMD functions quite broadly in the regulation of gene expression (Kervestin and Jacobson, 2012; Lelivelt and Culbertson, 1999; Medghalchi et al., 2001; Rehwinkel et al., 2005).

The central molecular players of NMD are the three conserved "up-frameshift" proteins Upf1, Upf2, and Upf3 which were originally isolated in genetic screens for increased frameshifting output (Culbertson et al., 1980; Leeds et al., 1992). Upf1 is the catalytic component of this complex, structurally organized into an N-terminal cysteine-histidine-rich (CH) domain that interacts with Upf2 and other proteins, and a RecA-like helicase core (SF1 family) that is responsible for RNA-dependent ATPase and helicase

activities (Chakrabarti et al., 2011; Chamieh et al., 2008; Czaplinski et al., 1995; Weng et al., 1998). In contrast, Upf2 and Upf3 are thought of primarily as interacting partners that regulate when and where Upf1 activity is implemented (Chamieh et al., 2008; Kervestin and Jacobson, 2012; Schoenberg and Maquat, 2012). Despite much molecular understanding of interactions between the Upf proteins and their binding partners, we know little about the earliest events that trigger NMD on a particular RNA. It has long been known that translation is required for NMD (Losson and Lacroute, 1979; Takeshita et al., 1984), and it stands to reason that events on the ribosome at the premature termination codons will be critical. It also has been shown that Upf1 migrates with polysomes (Atkin et al., 1995), and more recently that it can bind ribosomal protein eS26 in isolation (Min et al., 2013), but the molecular details of interactions with intact ribosomes remain poorly characterized.

Using purified components and a high throughput sequencing-based Fe-BABE directed hydroxyl radical probing method (HTS-BABE) we have characterized the interaction between Upf1 and the 80S ribosome. We identify the L1 stalk of the 25S rRNA as the binding site for Upf1 on 80S ribosomes, and show that this interaction is dependent on the 1C domain of Upf1, a small alpha-helical insertion in the RecA helicase core. Our Fe-BABE probing data and structural modeling with high-resolution structures of the human 80S ribosome and Upf1 allow us to propose that the interaction between Upf1 and the L1 stalk requires an open configuration, such as that populated by ribosomes with tRNAs bound in a classical state. Taken together our results shed light on this critical interaction between Upf1 and the ribosome, and suggest a function for Upf1 in stabilizing the classical state of the ribosome during translation.

RESULTS

Upf1 directly interacts with 80S ribosome via 1C domain

Several previous studies have shown that Upf1 sediments with 80S ribosomes in polysome gradients (Atkin et al., 1995) from cell lysate or binds ribosomal proteins in isolation (Min et al., 2013). To better understand the molecular details of the interaction between Upf1 and 80S, we used *in vitro* purified proteins (including full-length Upf1 and many variants) (Figure 1A and S1A) and ribosomal subunits to perform ribosome-pelleting assays (Figure 1B-F and S1B). In these experiments, ribosomal subunits and Upf1 were spun through a sucrose cushion and the pellet analyzed by Western blot to determine the extent of ribosome-bound Upf1. In an initial reaction, full-length Upf1 efficiently pelleted with 80S yeast ribosomes but not bacterial 70S ribosomes (Figure 1B, lane 3 compared to lane 1). These observations argue that the interaction between Upf1 and the eukaryotic ribosome is not simply due to the general RNA-binding activity of Upf1. We also note that this ribosome binding activity was readily competed by the addition of high amounts of polyU in the binding reaction (Figure 1B, lane 4), consistent with the fact that Upf1 is known to generally bind RNA and that its ATPase activity is stimulated by RNA (Czaplinski et al., 1995).

Because the ATPase function of Upf1 is known to be critical for NMD in cells, we next evaluated the nucleotide dependence of the Upf1:80S interaction. Although Upf1 interacts with the ribosome under all nucleotide conditions (ATP, ADP, ADPNP, or no nucleotide), binding was most effective in the presence of ADP or no nucleotide relative to ADPNP or ATP (Figure 1C). These data correlate nicely with reported binding constants for the Upf1:polyU interaction under similar conditions (Chakrabarti et al.,

2011), and further suggest that Upf1 might interact with both rRNA and mRNA through similar surfaces. We next asked whether known mutations in Upf1 that perturb ATP binding (K436Q), ATP hydrolysis (DE572AA), or RNA binding (RR793AA) impact the interaction with the ribosome. We found that Upf1 ATPase mutants DE572AA and K436Q efficiently pellet with the ribosome, while the RNA binding mutant (RR793AA) pelleted with reduced efficiency (Figure 1D). Each of these variant proteins was as readily competed from the ribosome by an excess of polyU RNA as the wild-type Upf1.

Previous reports suggested a role for the N-terminal CH domain of Upf1 in interaction with the ribosome (Min et al., 2013). To test this directly, we purified a variant of Upf1 lacking this domain as well as a more severe truncation containing only the helicase core of Upf1 (Figure 1A; Upf1 ΔCH and Upf1 NΔ373 respectively). Ribosome pelleting assays with these constructs showed that the CH domain was not critical for ribosome binding *in vitro* (Figure 1E). Interestingly, even the Upf1 helicase core (Upf1 NΔ373) was efficiently pelleted, although binding by this variant protein was less efficiently competed by polyU (likely because much of the mRNA binding surface area contained in the CH and 1B domains has been removed).

We next focused on the 1C domain of Upf1, a small, conserved insertion in the first RecA domain of Upf1 that is comprised of a three alpha-helix bundle that positions two positively charged patches on the surface of Upf1. One surface points away from the previously defined mRNA binding channel (Helix 1,2) (Figure 1F, cyan residues) and another points toward the mRNA binding channel (Helix 3) (Figure 1F, light purple residues). We first generated a Upf1 variant where the 1C domain (residues V494-K546) was substituted with a stretch of twelve alanine residues (Upf1 1C-12Ala) and found that

Upf1 binding to the ribosome was very nearly abolished (Figure 1F, lane 3). More selective alanine substitution of several positively charged patches on either of two sides of the helical bundle revealed Helix 3 as the more critical surface for ribosome binding (Figure 1F, lane 5). This same surface has previously been shown in X-ray structures to be involved in mRNA binding (Chakrabarti et al., 2011). In fact, substitution of the 1C domain with 12 alanines greatly weakens the mRNA binding ability of Upf1 as evidenced by our data (Figure S1C) and others (Cheng et al., 2007). Taken together, our results indicate that Upf1 binds ribosomal rRNA through the positively charged 1C domain and are consistent with a dual-binding mode for this domain that could allow Upf1 to partition between mRNA and the ribosome during NMD.

Development of HTS-BABE method using eIF5A

In order to more precisely define the site of interaction between Upf1 and 80S ribosomes, we pursued a hydroxyl radical probing approach using site-specific Fe-BABE labeling. Radical probing/footprinting has been used for decades to identify protein binding sites on the ribosome (Powers and Noller, 1995). This approach works by using Fenton chemistry to create hydroxyl radicals (in the presence of Fe(II), hydrogen peroxide, and ascorbic acid) that directly cleave nucleic acid backbone (Dixon et al., 1991). To add an additional layer of control to this reaction, Fe-BABE (Fe(II)-bromoacetamidobenzyl-EDTA) molecules can be site-specifically conjugated at cysteine residues on a protein of interest to spatially localize hydroxyl radical generation (Rana and Meares, 1991). In this way, hydroxyl radicals are created at the Fe-BABE modification site and can diffuse as far as 40 Å to cleave rRNA (Joseph et al., 1997). This directed-probing has been used successfully to map protein binding sites on the ribosome

(Culver et al., 1999; Heilek and Noller, 1996; Wilson and Noller, 1998), as well as conformational rearrangements that occur during translation (He and Green, 2010). The cleaved rRNA bases are detected by reverse transcriptase stops in a primer extension reaction when comparing a sample that has been labeled with Fe-BABE at a particular location to one lacking this modification. Although this method has proven powerful over the years, one limitation of the technique is that it requires many primer extension reactions and gels to identify cleavage sites across the entire ribosomal rRNA, which is costly and extremely time consuming.

To improve this method, we adapted Fe-BABE-mediated rRNA cleavage detection to high-throughput sequencing technology (HTS-BABE, Figure 2A) as recently described for a solution-based experiment aimed at defining the accessible surface area of the *E. coli* 16S rRNA (Kielpinski and Vinther, 2014). Our method uses directed hydroxyl radicals created at a site-specific Fe-BABE moiety, rather than those created in solution, to allow for high resolution mapping of protein binding sites on the ribosome. We chose to test our method using eIF5A, a small protein known to bind the ribosomal E site to stimulate global translation elongation and termination (Schuller et al., 2017). eIF5A was an ideal candidate to test our method because Fe-BABE probing has previously been used to map its specific ribosome binding site (Gutierrez et al., 2013), and a recently determined high-resolution cryoEM structure verified this initial positioning (Schmidt et al., 2016).

We expressed and isolated from *S. cerevisiae* two previously characterized site-specific cysteine variants of eIF5A (K48C and T126C) and a "scrubbed" version of eIF5A where all cysteines were removed (native C23A and C39T mutated; called eIF5A

Δcys) (Figure S2A). These proteins were modified with Fe-BABE and subjected to our HTS-BABE approach (Figure 2A). The method begins by incubating ribosomal subunits and Fe-BABE labeled eIF5A proteins in the presence of ascorbic acid and hydrogen peroxide to initiate Fenton chemistry and hydroxyl radical cleavage. The reactions are then quenched by addition of thiourea, and the ribosomal RNA from each reaction with Fe-BABE labeled protein or the eIF5A Δcys control is isolated. Subsequently, this cleaved rRNA is reverse transcribed using a random nonamer (SN_8)-containing oligonucleotide and superscript III reverse transcriptase to reveal cleaved bases as reverse transcriptase stops. These cDNA fragments are then subjected to a 3′ linker ligation and subsequent amplification using barcoded primers for Illumina sequencing (see Methods and Materials for details).

To identify the rRNA cleavage sites, we mapped our sequencing reads to each of the *S. cerevisiae* ribosomal RNAs and compared the number of RT stops at each position for each Fe-BABE variant with the eIF5A Δcys control. For eIF5A K48C, we found strongly enhanced rRNA cleavage (outliers from the MA plot) at 8 positions in the 25S rRNA compared to eIF5A Δcys (Figure 2B and S2B). For eIF5A T126C, we found enhanced cleavage at 6 nucleotides of the 25S rRNA and 6 nucleotides in the 18S rRNA (Figure 2C and S2C). Many of these sites overlap with those identified in previous Fe-BABE probing experiments evaluated with primer extension gels (Gutierrez et al., 2013); we also confirmed these cleavage sites using primer extension gel analysis (Figure 2D). Importantly, these locations mapped onto the 80S ribosome structure nicely clustered within 30 Å of the modified cysteine in the context of the previously solved cryoEM structure (Figure 2E and S2D-E) (Schmidt et al., 2016). These experiments performed

with eIF5A establish that this method can readily identify Fe-BABE hydroxyl radical cleavage locations without the need for primer extension gels. Moreover, with proper multiplexing of samples, it is possible to obtain data for many Fe-BABE variants in a single sequencing lane.

HTS-BABE reveals Upf1 binding to 25S ribosomal RNA

We next performed a similar HTS-BABE experiment with Upf1 (and 14 distinct site-specific cysteine variants) and the 80S ribosome (Figure 3A and S3A). As the Upf1 CH domain is cysteine-rich but dispensable for ribosome binding (Figure 1E), we chose a Upf1 ΔCH parent construct for this experiment (residues 221-851) (Figure S3B). While this truncated Upf1 construct contained 9 natural cysteines, the crystallographic structure suggested that 5 were not adequately surface-exposed for Fe-BABE modification and were thus left as is. We determined that the remaining 4 cysteines could be removed to create soluble, non-aggregated Upf1 (C709A, C777A, C833A, C845A; called Upf1 Δ4cys) that efficiently pelleted with 80S ribosomes (data not shown). This parent construct was then mutated to create site-specific cysteine variants that span the entire surface of Upf1, with several localized on or near the ribosome-binding 1C domain that we identified above (Fig. 1F). These Upf1 variants were expressed and purified from *E. coli* (see Materials and methods), labeled with Fe-BABE, and subjected to the same HTS-BABE method developed for eIF5A.

Our HTS-BABE analysis for these site-specific Upf1 Fe-BABE variants revealed two variants with sites of strongly enhanced rRNA cleavage compared to the Upf1 Δ4cys control. Upf1 S335C showed enhanced cleavage (outliers from the MA plot) at 25S

C2496 and U2497 (19 and 15-fold respectively) (Figure 3B and S3C) and T337C showed enhanced cleavage at 25S A2485 (23-fold) (Figure 3C and S3D). Both variants are positioned on the 1B domain of Upf1 that is spatially located adjacent to the ribosome-binding 1C domain (Figure 3D) (i.e. the key residues identified as critical to ribosome binding) (Figure 1F). Traditional primer extension analysis for Upf1 S335C (Figure 3E) and T337C (Figure 3F) confirmed these strong rRNA cleavage locations, and also identified other weaker cleavage sites near this region of the 25S rRNA for these variants and several others.

Upf1 1C domain is required for interactions with L1 stalk

The cleavage sites identified by Upf1 Fe-BABE probing map to the 25S rRNA L1 stalk (Figure 4A). This region of the ribosome has been shown to undergo large-scale movements throughout translation elongation in *E. coli* (Mohan and Noller, 2017). In classical-state ribosomes, where the E site is empty, the L1 stalk adopts an "open" conformation, oriented away from the ribosome core. After peptidyl-transfer, the deacylated tRNA (having just transferred the peptidyl group to the A-site tRNA) adopts a hybrid P/E state, and the L1 stalk moves inward to make interactions with the tRNA elbow in a "closed" state. Next the L1 stalk moves through two "intermediate-closed" positions as the tRNA moves through a pe/E chimeric hybrid state, and finally into a classical E/E state. The stalk maintains interactions with the tRNA elbow in these intermediate states, so release of the E-site tRNA likely occurs after the stalk fully opens again.

As the L1 region of the ribosome is RNA-rich and flexible, we wondered whether this binding site might simply reflect the RNA binding propensity of Upf1. Importantly,

however, our HTS-BABE analysis with these 14 Upf1 variants did not show significant, reproducible cleavage on other long extended helices on the ribosome such as 18S Helix 16 and expansion segment 6 (Figure S4A). While we did find modestly enriched cleavage on one long RNA helix (18S U649; Figure S4B) for several variants, and another variant yielded modestly enriched cleavage on the 18S Helix 21 expansion (Figure S3C), there was no evidence for cleavage at these sites using traditional primer extension analysis (Figure S4C-D). Taken together, we suggest that these sites are false-positives that emerge from our HTS-BABE experiment. We argue that the robust and reproducible rRNA cleavage that we observe at the L1 stalk of yeast 80S ribosomes using both HTS-BABE and traditional primer extension analysis is sufficient to define a binding site for Upf1 on the eukaryotic ribosome.

To further validate the Upf1 binding site, we asked whether the L1 stalk cleavages were dependent on the 1C domain of Upf1. We used the Upf1 S335C variant (which yielded a strong signal at bases C2496-U2498 of the L1 stalk along with other, weaker cleavage sites in the area) to probe the necessity of the 1C domain by constructing an S335C 1C-12Ala construct (1C domain replaced with twelve alanine residues) and performing the same Fe-BABE probing and primer extension analysis. This experiment revealed a loss of cleavage events in the L1 stalk region, suggesting that the 1C domain is required for the Upf1 interaction with the ribosome that we define by HTS-BABE (Figure 4B).

Model for Upf1 interaction with ribosomal L1 stalk

With the information provided by our HTS-BABE analysis, we can propose a preliminary model for how Upf1 may interact with the L1 stalk by rigid-body docking of the Upf1 structure on an 80S structure. First, however, as there is no current structural data for the complete L1 stalk in the yeast ribosome, we had to model Upf1 onto the structure of the human 80S ribosome where the L1 stalk is more defined (PDB 4UG0) (Khatter et al., 2015). Importantly, in this human structure, the L1 stalk adopts an "intermediate-closed" configuration in the presence of a classical-state E-site tRNA, though we anticipate that the human ribosome, like the *E. coli* ribosome, can also adopt an "open" configuration where L1 swings out away from the ribosome (and the E site is empty). In the "intermediate-closed" state, we were unable to model a Upf1 molecule bound to the L1 stalk without steric clash in the E site of the ribosome (data not shown), suggesting that ribosomes may have to adopt the open state in order to permit Upf1 binding.

As there are no structures of eukaryotic ribosomes with an open and resolved L1 stalk, we modeled the "open" state by moving the L1 stalk to align with the open position that has been observed in the *E. coli* 70S ribosome structure; the L1 stalk has been shown to move by as much as 38 Å as it transitions between the fully "open" and the "intermediate-closed" state (Mohan and Noller, 2017). We therefore moved the human L1 stalk by approximately 30 Å (Figure 4C) using a previous *E. coli* 70S structure where the L1 stalk is found in an open configuration (PDB 4V9D) (Dunkle et al., 2011) as a guide. This open state created enough space between the L1 stalk and ribosomal E site to easily proceed with a rigid-body placement of Upf1.

In this open state, we could model Upf1 bound to the L1 stalk by rigidly positioning the crystal structure for yeast Upf1 (PDB 2XZL; residues 221-851) (Chakrabarti et al., 2011) and using our observed rRNA cleavage sites to guide the positioning (Figure 4D-F). The Upf1 T337C variant cleaved A2484-A2486 of yeast 25S rRNA which corresponds to A4046-A4048 of human 28S rRNA (colored magenta); the S335C variant cleaved C2496-U2498 of yeast 25S rRNA which corresponds to U4058-U4060 human 28S rRNA (colored cyan). This conformation places Upf1 at the edge of the ribosomal E site, with room for the E site tRNA to remain accommodated (Figure 4E), and places the 1C domain where it can make contacts along Helix 76 of the L1 stalk (Figure 4E-F). In this position, Upf1 is seen to be in close proximity to ribosomal protein eS26 at the mRNA exit channel consistent with previous reports of interaction between the CH domain of Upf1 and eS26 in *S. cerevisiae* (Min et al., 2013). While we removed the CH domain from our modeling exercise (as it was not essential for ribosome binding or used in our probing experiments), it is possible that the CH domain could interact with eS26 in this position after a large conformational change such as that reported upon Upf2 binding (Chakrabarti et al., 2011; Clerici et al., 2009).

Upf1 does not affect *in vitro* reconstituted elongation, termination, or recycling reactions

Given the binding site of Upf1 on the L1 stalk, and that large-scale movements of this region correlate with the ribosome functional state, we wondered if Upf1 could affect translation directly through interactions with L1 stalk. To test this possibility, we added purified Upf1 (the Upf1 ΔCH variant used for binding site identification) to several well-

established assays in a fully purified *in vitro* reconstituted system (Eyler and Green, 2011; Schuller et al., 2017; Shoemaker et al., 2010; Shoemaker and Green, 2011). We first added Upf1 to initiated 80S ribosomes and performed an elongation reaction where we monitored penta-peptide (MFFFK) synthesis over time (in the presence of the appropriate aminoacyl-tRNAs and the required yeast elongation factors eEF1A, eEF2, and eEF3). We see no effect of Upf1 on MFFFK synthesis rate (Figure 5A), suggesting that under these conditions, L1 stalk binding by Upf1 has no strong influence on translation elongation. Next we tested Upf1 in an *in vitro* termination reaction, monitoring peptide release by the hydrolytic activity of eRF1. As for the elongation reaction, Upf1 addition had no effect on this assay (Figure 5B). Finally, we tested the effect of Upf1 on ribosome subunit recycling catalyzed by Rli1 (ABCE1 in mammals) and eRF1 and, again, observed no effect (Figure 5C). While these assays were repeated multiple times, using multiple different conditions to "sensitize" the assay, across several preparations of Upf1 including constructs that contained the N-terminal CH domain or were ATPase inactive (data not shown), we were unable to observe any significant biochemical effects. Taken together, these results do not support a direct role for Upf1 in ribosome function, though we recognize that the *in vitro* reconstituted system may lack components important for Upf1 activity to be manifested.

DISCUSSION

The nonsense-mediated decay pathway in eukaryotes that selectively degrades PTC-containing mRNAs depends on three centrally important Upf proteins (Upf1, Upf2, and Upf3). Although much work has been done to shed light on the interactions between the Upfs and other proteins, as well as to characterize the factors involved in mRNA degradation, many questions remain as to how NMD is initiated on the ribosome. Upf1 is the central, catalytic component of the NMD pathway and has been shown to interact with 80S ribosomes, likely in an early step of NMD. However, the molecular details and functional consequence of this interaction remain poorly characterized.

Here, using *in vitro* purified components and a new, sequencing-coupled Fe-BABE directed hydroxyl radical approach, we characterize the binding interaction between Upf1 and the 80S ribosome. We show that the Upf1 1C domain, a small alpha-helical insertion in the RecA-like helicase core, is critical for ribosome (Figure 1) and mRNA binding (Figure S1C), suggestive of a partitioning event mediated by this domain between the mRNA and ribosome during NMD. To identify the binding site for Upf1 on the 80S ribosome, we developed a high throughput sequencing-based Fe-BABE directed hydroxyl radical probing method (HTS-BABE) (Figure 2), and identified the L1 stalk of the 25S rRNA as the binding site of Upf1 (Figure 3). We subsequently verified this interaction to be dependent on the 1C domain of Upf1 and assembled a model of Upf1 binding to the L1 stalk on human 80S ribosomes (Figure 4).

Despite the robust nature of the interaction that we define, the molecular details of the Upf1:80S interaction were somewhat surprising to us. The NMD pathway must depend on recognition of a premature termination codon, an event that occurs in the

ribosomal A site, while our data suggest that Upf1 binds on the other side of the ribosome in the region proximal to the E site. Although other published data have suggested that Upf1 binding occurs in this region near eS26 (Min et al., 2013), it remains unclear how this binding site is connected to proper initiation of NMD. We speculate that Upf1 binding to the L1 stalk might have long-range effects on the ribosome that impact A site reactivity as observed for a number of other translation factors (eEF3, eIF5A, EFP, and EttA) (Andersen et al., 2006; Blaha et al., 2009; Boel et al., 2014; Chen et al., 2014; Doerfel et al., 2013; Gutierrez et al., 2013; Schmidt et al., 2016; Schuller et al., 2017; Triana-Alonso et al., 1995; Ude et al., 2013). This action is possible given the highly dynamic nature of the L1 stalk as it moves from a "closed" position with hybrid-state tRNAs after peptidyl-transfer, to an "open" conformation with a vacant E site. As translation termination and recycling both occur at termination codons (a hallmark of NMD), and since the ribosome would be anticipated to be in a classical state during these events (Brown et al., 2015; Matheisl et al., 2015; Preis et al., 2014; Shao et al., 2016), it is possible that Upf1 could impact one of these functions through long-range allosteric effects (Nierhaus, 1990). Whatever the mechanism of action for Upf1 might be, our identification of the 1C domain as critical for the Upf1:80S interaction is strongly supported by previous work indicating that this domain is essential for NMD (Cheng et al., 2007).

While our attempts at observing a direct effect for Upf1 in translation elongation, termination, or recycling using an established *in vitro* reconstituted system were unsuccessful (Figure 5), it is possible that our reactions lack some key component necessary to reveal Upf1 activity. Importantly, Upf1 did not have any effect on

translation termination, though Upf1 is believed to interact with eRF1 based on co-immunoprecipitation experiments (Ivanov et al., 2008) and direct inhibition of Upf1's ATPase function by eRF1 (Czaplinski et al., 1998). Additionally, while other molecules such as cycloheximide or eIF5A bind inside the ribosomal E site to affect A-site reactivity (Budkevich et al., 2011; Garreau de Loubresse et al., 2014; Gutierrez et al., 2013; Schmidt et al., 2016; Schneider-Poetsch et al., 2010; Schuller et al., 2017), the interaction between Upf1 and the L1 stalk likely occurs outside of the ribosomal E site, and thus may not directly affect L1 movement or ribosome function. For example, it is possible that the Upf1-ribosome interaction is critical for interactions with additional factors not included in our *in vitro* reconstituted system.

We anticipate that our HTS-BABE approach will enhance structural determination by radical probing methods, or be used in conjunction with other methods such as x-ray crystallography and cryoEM. While the resolution of this method is not atomic, it could be used to cross-validate cryoEM data when the identity of certain volumes is unknown, to help orient a molecule if sufficient resolution cannot be obtained, or be used when molecules are flexible or otherwise suboptimal for structural determination. In the case of Upf1, the eukaryotic ribosomal L1 stalk remains poorly characterized by these other structural methods, likely because of its inherent flexibility, and it may be challenging to image unless it can be trapped in a stable conformation. Future pursuits to stabilize and determine the complete Upf1:80S structure will be required to support our understanding of how Upf1 might initiate NMD at a terminating ribosome.

MATERIALS AND METHODS

Full-length Upf1 expression and purification

The Upf1 gene from *S. cerevisiae* (NAM7) was cloned into pYES-DEST52 (Invitrogen) with TEV protease-cleavable N-terminal His_7 and C-terminal maltose binding protein (MBP) tags for affinity purification. All variants of Upf1 (point mutants and truncations) were prepared by sub-cloning from this parent plasmid using Topo Directional cloning (Invitrogen) or QuickChange mutagenesis (Stratagene). The pYES-DEST52 plasmid contains a *GAL1* inducible promoter for expression of Upf1 in *S. cerevisiae*. All full-length Upf1 proteins were expressed in *S. cerevisiae* JC287 (MATa, *ade2-1, his3, leu2, trp1, ura3, pep4::HIS3, prb1::HIS3, prc1::HIS3*) (provided by Dr. Jeff Coller) with 2% galactose. Cell pellets were harvested, flash frozen in A500 buffer (25 mM Tris, pH 7.5, 500 mM NaCl, 10% glycerol, 1 mM MgCl2, 1 μM ZnCl2, and 5 mM 2-mercaptoethanol), and lysed in a liquid nitrogen Freezer/Mill (SPEX Sample Prep, LLC). Cell lysates were purified over an amylose column (NEB) followed by a HisTrapFF column (GE Healthcare) to yield pure full-length Upf1 protein. Protein fractions were pooled, concentrated, exchanged into buffer A150 (same as A500 but with 150 mM NaCl and 1 mM DTT instead of 2-mercaptoethanol) and stored at -80^{o}C. In our pelleting assays, the His and MBP tags were not cleaved from Upf1 as initial experiments showed they did not affect function (Figure S1B).

Ribosomal subunit purification and pelleting assay

Ribosomal subunits from *S. cerevisiae* were purified as previously reported (Eyler and Green, 2011), or from *E. coli* as previously reported (Youngman et al., 2004).

Ribosome pelleting assays were conducted by incubating ribosomal subunits with Upf1 proteins in reactions containing 100nM ribosomal subunits, 100 nM Upf1, 1 mM nucleotide, and 0.2 mg/ml polyU (Sigma) in 1X Buffer E (20 mM Tris, pH 7.5, 100 mM KOAc pH 7.5, 2.5 mM $Mg(OAc)_2$, 0.25 mM spermidine, and 2 mM DTT). Reactions were pelleted over sucrose cushions containing 1.1 M sucrose in Buffer E using a Beckman MLA-130 rotor at 75,000 rpm for 1 hour at 4°C. Pelleted ribosomes were resuspended in Buffer E and run on SDS-PAGE gels for Western blotting using an anti-MBP antibody (NEB) to detect Upf1.

Upf1 221-851 expression and purification

Residues 221-851 of the Upf1 gene were cloned into a plasmid containing a TEV protease-cleavable N-terminal His_6-MBP tag (pDESTHisMBP). Each construct was designed to include an N-terminal HA-tag in order to observe ribosome pelleting by Western blot analysis. This Upf1 construct was transformed into Rosetta 2 (DE3) pLysS competent cells (Novagen) and expressed overnight in terrific broth (RPI) in the presence of 500 μM IPTG at 16°C. Cells were lysed by French press (Thermo) in A500 buffer and purified using amylose resin (NEB). Fractions containing Upf1 protein were then pooled, and incubated with His-tagged TEV protease overnight to remove the His_6-MBP tag on Upf1. After cleavage, proteins were further purified over Ni-NTA resin (Qiagen) to remove both the free His_6-MBP and His_6-tagged TEV protease, and a Superose 6 10/300 GL column to ensure the protein was not aggregated. Purified Upf1 was stored in A150 buffer at -80°C for subsequent assays.

Fe-BABE labeling and rRNA cleavage

eIF5A site-specific cysteine variants were labeled with Fe-BABE by first dialyzing 40 μL of 40 μM protein into eIF5A modification buffer (30 mM Hepes, pH 7.5, and 500 mM KCl) for 5.5 hours at 4^{o}C. After dialysis, each eIF5A variant was labeled by incubation with 2 mM Fe-BABE (Dojindo Molecular Technologies) at 30^{o}C for 30 minutes. After labeling, the proteins were concentrated and buffer exchanged into Fe-eIF5A storage buffer (20 mM Hepes, pH 7.5, 100 mM KCl, and 3 mM $MgCl_2$), and stored at -80^{o}C.

Upf1 site-specific cysteine variants were labeled using a similar protocol with a few modifications for Upf1 protein stability. 40 μL of 40 μM Upf1 protein was dialyzed into Upf1 modification buffer (30 mM Hepes, pH 7.5, 10% glycerol, and 500 mM KCl) for 5.5 hours at 4^{o}C. After dialysis, each Upf1 variant was labeled by incubation with 2 mM Fe-BABE (Dojindo Molecular Technologies) at 30^{o}C for 30 minutes, or in one circumstance at 4^{o}C overnight (Figure 4B labeled as "o/n"). After labeling, the proteins were concentrated and buffer exchanged into Fe-Upf1 storage buffer (20 mM Hepes, pH 7.5, 150 mM KCl, 10% glycerol, and 3 mM $MgCl_2$), and stored at -80^{o}C.

rRNA cleavage was performed as a 20 μL reaction containing 500 nM 80S ribosomal subunits (purified as described above), 500 nM Fe-BABE-labeled protein (eIF5A or Upf1) in cleavage buffer (20 mM Hepes, pH 7.5, 100 mM KCl, and 3 mM $MgCl_2$). Reactions were incubated on ice for 15 minutes to allow factor binding. Then we added 1 μL of 100 mM ascorbic Acid, and 1 μL fresh 0.5% hydrogen peroxide simultaneously to initiate radical generation. Pipetting each 1 μL reactant on the wall of a microcentrifuge tube, and then quickly spinning the contents allowed for simultaneous

mixing. The reaction continued on ice for 15 minutes and was then quenched by addition of 1 μL of 100 mM thiourea and 300 μL of rRNA extraction buffer (0.3 M NaOAc). The rRNA was then extracted by acid phenol, phenol:chloroform (5:1), and subsequently ethanol precipitated and resuspended in nuclease-free H_2O.

HTS-BABE method and analysis

The HTS-BABE method contains several steps: 1) Reverse transcription of the cleaved rRNA, 2) Cleanup by RNAClean XP beads (Agencourt), 3) 3′ end ligation by Circligase (Epicentre), 4) Cleanup by Ampure XP beads (Agencourt), 5) PCR and gel extraction of products, and 6) Illumina sequencing. As previously mentioned, our method is based on one recently published (Kielpinski and Vinther, 2014) with several modifications we have described here.

In step 1, the cleaved rRNA is reverse transcribed (RT) using a random nonamer (SN_8)-oligo (oAS544). 1 μg of cleaved rRNA was incubated with 20 μM oAS544 in 1X RT-Mg buffer (50 mM Tris-Cl, pH 8.6, 60 mM NaCl, 10 mM DTT) at 65^{o}C for 5 minutes and then moved to ice to allow annealing of the oligo on the rRNA. Then 5 μL of this annealed-mixture was added to a mixture containing 2.5 μL 5X dNTP mix (1.7 mM of each dNTP), 2.5 μL 1X RT+Mg buffer (50 mM Tris-Cl, pH 8.6, 60 mM NaCl, 10 mM DTT, 6 mM $MgCl_2$), and 2.5 μL RVT mix (0.25 μL 10X RT-Mg buffer, 0.25 μL SSIII reverse transcriptase (Thermo), and 2 μL H_2O) to yield a final volume of 12.5 μL. This reaction was then reverse transcribed by incubating as follows: 25^{o}C – 10 min, 50^{o}C – 70 min, 60^{o}C – 10 min, 4^{o}C – indefinite. After the reaction completed, we incubated at 70^{o}C for 15 min to denature the reverse transcriptase and then added 25 μL of 1X RT+Mg

buffer to make a final volume of 37.5 µL. The reaction was then subjected to RNaseH digestion by adding 1 µL of RNAseH (NEB) and incubating at 37°C for 20 min to digest the rRNA.

In step 2, the resulting reverse transcription product is purified by RNAClean XP beads. We added 67.5 µL of beads to each reaction, and incubated at room temperature for 30 minutes, mixing the components every 10 minutes. Then the supernatant was removed after placing the reactions on a magnetic stand for 5 minutes. We subsequently washed with 200 µL of 70% EtOH two times, and eluted with 40 µL of 5 mM sodium citrate, pH 6.0. To elute, the beads were mixed with buffer and incubated at 37°C for 10 minutes, then placed on the magnetic stand and the eluant removed.

In step 3, the cDNA product is ligated with a 3′ linker for future sequencing steps. In this reaction, we mixed 3 µL of each cDNA product with a 7 µL ligation mix containing 1 µL 10X Circligase buffer, 0.5 µL 1 mM ATP, 0.5 µL 50 mM $MnCl_2$, 2 µL 50% PEG 6000, 2 µL 5 M betaine, 0.5 µL of 100 µM oAS545 linker, and 0.5 µL Circligase enzyme. The resulting 10 µL reaction was incubated as follows: 60°C – 120 min, 68°C – 60 min, 80°C – 10 min, 4°C – indefinite. Then we added 10 µL H_2O for a final volume of 20 µL to proceed.

In step 4, this ligated cDNA product is purified using Ampure XP beads. The procedure is similar to step 2 with the following exceptions: 1) We used 36 µL of beads to clean the 20 µL sample, and 2) We eluted using 16 µL of H_2O.

In step 5, the purified cDNA is amplified using PCR primers containing adaptor sequences for Illumina sequencing and barcodes for sample multiplexing. Our PCR reactions contained 6 µL of ligated cDNA product, 3.6 µL of 10 µM Index oligo

(oASBar01-30), 3.6 µL of 10 µM forward oligo oBZ287, 12 µL 5X Phusion HF Buffer, 4.8 µL 2.5 mM dNTPs, 28.8 µL H_2O, and 1.2 µL Phusion DNA polymerase. Our amplification protocol was as follows: 98°C – 3 min, (98°C – 80 sec, 64°C – 15 sec, 72°C – 30 sec) repeated 4 times, (98°C – 80 sec, 72°C – 45 sec) repeated 10 times, 72°C – 5 min, 4°C – indefinite. The PCR reactions were then run on an 8% TBE native gel at 200V for 28min and stained with SYBR gold. The resulting products > 150 bp were cut from the gel and extracted overnight in 0.3 M NaOAc with 1 mM ETDA at room temperature. We intentionally cut >150 bp to avoid sequencing the amplified product that results from ligation of the unextended RT oligo directly to the 3′ linker (and therefore contains no insert cDNA). The following day, the extracted DNA was precipitated with isopropanol and quantified using an Agilent 2100 Bioanalyzer before sequencing.

The amplified products were subjected to 50bp single-end sequencing on an Illumina HiSeq 2500 and subjected to our analysis pipeline. The reads are first trimmed of the 3′ adaptor using the skewer package (Jiang et al., 2014). The N_7 randomized nucleotides from our ligated adaptor were trimmed from read 5′ ends using fastx_trimmer (http://hannonlab.cshl.edu/fastx_toolkit/index.html by Hannon Lab). The read 5′ end now corresponds to the 5′ nucleotide of the 3′ product resulting from radical cleavage. Next, the reads were mapped to the yeast ribosomal rRNA (including 25S, 18S, 5.8S, and 5S) using STAR (Dobin et al., 2013) to provide a normalized (reads per million, rpm) read 5′ end count at each rRNA nucleotide.

To analyze our data, we compared the rpm at each rRNA nucleotide for each site-specific cysteine variant with the eIF5A Δcys or Upf1 Δ4cys controls. We simply divided the rpm at each nucleotide position across the ribosome to generate a fold-change of

reads in the labeled variant compared to the control, and thereby to identify enhanced cleavage sites caused by the Fe-BABE label at a particular position.

Sequencing data and mapped read counts have been deposited in GEO with accession number GSE104072.

Primer extension for rRNA cleavage analysis

To identify rRNA cleavage sites by primer extension analysis, first the cleaved rRNA must be reverse-transcribed with a radioactive site-specific oligo complementary to the rRNA sequence. Approximately 1 μg of cleaved rRNA is first mixed in a 5 μL reaction with a 32-P end-labeled oligonucleotide (200 nM) in 1X RT-Mg buffer and annealed by incubating at 60^{o}C for 5 minutes and then moved to ice. 2 μL of this annealed-mixture was then added to a mixture containing 1 μL 5X dNTP mix (1.7 mM of each dNTP), 1 μL 1X RT+Mg buffer, and 1 μL RVT mix (0.25 μL 10X RT-Mg buffer, 0.25 μL AMV (Roche), and 2 μL H_2O) to yield a final volume of 5 μL. This reaction was then incubated at 42^{o}C for 60 minutes for reverse transcription to occur. For sequencing ladders, the reactions contained 2 μL of the annealed-mixture, 1.3 μL 5X dNTP mix, 0.7 μL 5X ddNTP (0.2 mM of one ddNTP), and 1 μL RVT mix. The resulting cDNA products were mixed with 2X formamide loading dye and run on 10% TBE-Urea gels at 60W for 2-4hrs depending on the region to be analyzed.

ATPase Assays

Purified full-length Upf1 constructs (200 nM) were mixed with 0.02 mg/ml polyU (for Upf1 and Upf1 ΔCH) or 0.2 mg/ml polyU (for Upf1 1C-12Ala). Reactions were

incubated at 26°C and time points quenched with 30% formic acid. Samples were spotted on PEI-Cellulose F TLC plates (EMD Millipore) and analyzed in 0.5 M KH_2PO_4 pH 3.5. TLC plates were developed using a Typhoon FLA 9500 Phosphorimaging system and quantified using ImageQuant TL (GE Healthcare Life Sciences).

***In vitro* reconstituted translation assays**

The elongation and termination assays were performed as previously described (Schuller et al., 2017) using MFFFKX-UAA initiation complexes or MFX-UAA pre-termination complexes with the addition of 10-20 μM Upf1 ΔCH.

The recycling assay was performed using MFKX-UAA pre-termination complexes in reaction with 1X Buffer E, 4 μM AGQ-eRF1, 2 μM Rli1, and 10 μM Upf1 ΔCH. 50 μM peptidyl-tRNA hydrolase (PTH) was added to the reactions to quantify dissociated complexes (PTH cannot access the peptidyl-tRNA unless released from the ribosome) (Shoemaker et al., 2010). Time points in both assays were quenched using 10% formic acid and run on electrophoretic TLC (Millipore).

All TLC plates were developed using a Typhoon FLA 9500 Phosphorimager system and quantified using ImageQuantTL (GE Healthcare Life Sciences). Time courses were fit to single exponential kinetics using Kaleidagraph (Synergy Software).

FIGURE LEGENDS

Figure 1. Upf1 binds 80S ribosomes via 1C domain

(A) Schematic of Upf1 domain structure with point mutants and domain truncation constructs used in ribosome pelleting assays. (B) Ribosome pelleting comparison of *S. cerevisiae* 80S ribosomes and *E. coli* 70S ribosomes. (C) Ribosome pelleting assay with different ATP analogs added. (D) Ribosome pelleting assay using Upf1 point mutants for ATP binding (K436Q), ATP hydrolysis (DE572AA), and weakened mRNA binding (RR793AA). (E) Ribosome pelleting assay using Upf1 N-terminal domain truncation constructs. (F) Ribosome pelleting assay using Upf1 construct with 1C domain replaced by a stretch of 12 alanines, or specific alanine mutants on each positively charged surface. Structure of Upf1 with 1C domain shown in red and mRNA in blue (PDB 2XZL). Positive residues on each side of the 1C domain are colored in cyan (Helix 1, 2) and light purple (Helix 3).

Figure 2. HTS-BABE method development and eIF5A test case

(A) Schematic of HTS-BABE workflow. (B) MA plot for eIF5A site-specific cysteine variant at position K48C and its corresponding rRNA cleavage sites identified using HTS-BABE. (C) Same as (B) for position T126. (D) Primer extension gels to probe sites identified by HTS-BABE method. (E) Structural model of eIF5A bound to 80S ribosome (PDB 5GAK) with Fe-BABE label locations (K48C – green, T126C – purple) and their corresponding cleavage sites as identified by HTS-BABE.

Figure 3. HTS-BABE reveals Upf1 binding to the ribosomal L1 stalk

(A) Pymol figure highlighting 14 Upf1 site-specific cysteine variants for Fe-BABE probing (PDB 2XZL; residues 221-851). Variants are colored in purple, and mRNA in blue for reference. (B) MA plot for Upf1 S335C position with top cleavage sites labeled. (C) Same as (B) for Upf1 T337C. (D) Pymol figure showing positions of S335C (cyan) and T337C (magenta) in relation to 1C domain and the helix 3 residues required for ribosome binding (shown in light purple). (E) Primer extension gel to probe cleavage sites along L1 stalk in S335C and other variants. (F) Same as (E) for an expanded set of Fe-BABE variants. Sample "S335C-Fe (Fig 4B)" was loaded as a gel control.

Figure 4. Upf1 1C domain is essential for L1 stalk binding

(A) Schematic of L1 stalk with unique identified cleavage sites from Upf1 variant S335C (cyan) and T337C (magenta). Cleavage sites shared between several variants are noted with black line (25S 2441-2443 and 2462-2464). (B) Primer extension gel to probe hydroxyl radical cleavage sites along L1 stalk in presence and absence of Upf1 1C domain. (C) Pymol model showing the movement of L1 stalk we employed to create an "open configuration." Open L1 stalk rRNA is shown in blue and L1 protein in green. The "intermediate-closed" L1 position (PDB 4UG0) is shown in grey for reference as well as the E-site tRNA in pink. (D) Pymol model for Upf1 binding to L1 stalk of human 80S ribosomes. (E) Close-up view of structural model highlighting Upf1 position in relation to L1 stalk and E-site tRNA, and cleaved residues by Upf1 S335C (cyan) and T337C (magenta). (F) Close-up view of Upf1 model their corresponding cleavage sites mapped to the L1 stalk rRNA.

Figure 5. Upf1 does not affect elongation, termination, or recycling in an *in vitro* reconstituted system

(A) *In vitro* elongation assay for MFFFK synthesis in the presence and absence of Upf1 ΔCH (residues 221-851). (B) *In vitro* termination assay monitoring peptidyl-release by eRF1:eRF3 in the presence and absence of Upf1 ΔCH. (C). *In vitro* recycling assay monitoring peptidyl-tRNA release by Rli1 in the presence and absence of Upf1 ΔCH.

SUPPLEMENTAL FIGURES & TABLES

Figure S1. Upf1 pellets with 80S ribosomes and the Upf1 1C domain affects both ribosome and mRNA binding

(A) Coomassie blue stained gel of Upf1 constructs containing an N-terminal His_7 and C-terminal MBP tag purified from yeast for pelleting assays. (B) Pelleting assay in the presence of TEV protease to cleave MBP and His_7 and test for non-specific pelleting by MBP. (C) ATPase assay for Upf1 FL, Upf1 ΔCH, and Upf1 1C-12Ala with varying amounts of polyU RNA.

Figure S2. HTS-BABE replicates for eIF5A and measurements of eIF5A Fe-BABE rRNA cleavage

(A) Coomassie blue stained gel of Flag-eIF5A constructs purified from yeast to be used in Fe-BABE cleavage assays. Gel includes flow through (FT), washes (W1-2) and elution (El) from FLAG immunoprecipitation. (B) MA plot for eIF5A K48C variant replicate in HTS-BABE method. Enriched cleavage sites are colored and labeled next to plot. (C) Same as (B) for eIF5A T126C variant. (D) Pymol figure of 80S ribosome bound to eIF5A (PDB 5GAK) with several identified rRNA cleavage sites for K48C variant mapped and measured for distance from Fe-BABE modification site. (E) Same as (D) for eIF5A T126C variant.

Figure S3. Upf1 E. coli purified ΔCH constructs pellets with 80S ribosomes dependent on 1C domain and HTS-BABE replicates for Upf1 S335C and T337C

(A) Coomassie blue stained gel of Upf1 site-specific cysteine variants used in Fe-BABE rRNA cleavage experiments. (B) Pelleting assay for E. coli purified Upf1 ΔCH and Upf1 ΔCH 1C-12Ala constructs. (C) MA plot for replicate in HTS-BABE experiment of Upf1 variant S335C. Enriched cleavage sites are labeled on each plot. In the case of S335C, several sites along Helix 16 were identified in HTS-BABE analysis, however later shown as false-positives (Figure S4D). (D) Same as (C) for Upf1 T337C variant.

Figure S4. Upf1 Fe-BABE rRNA cleavage analysis of other surface exposed long rRNA helices

(A) Pymol representation of *S. cerevisiae* 80S ribosome (PDB 4V88) to show solvent accessible RNA surface. 60S rRNA is colored in blue, 40S rRNA is colored in yellow, and ribosomal proteins are colored in gray. Helix 16 and expansion segment 6 are highlighted on structures as examples of long rRNA helices tested for rRNA cleavage with our Upf1 Fe-BABE variants. (B) MA plot indicating minor cleavage event at 18S U649 observed in several Upf1 variants but later shown to be a false-positive. (C) Primer extension analysis of Fe-BABE rRNA cleavage sites for several Upf1 variants along the 18S expansion segment 6 (ES6). (D) Primer extension analysis of Fe-BABE rRNA cleavage sites for several Upf1 variants along the 18S Helix 16. In (C) and (D), sample "S335C-Fe (Fig 4B)" was added as a gel control.

Table S1. Oligonucleotides used in this study

REFERENCES

Andersen, C.B., Becker, T., Blau, M., Anand, M., Halic, M., Balar, B., Mielke, T., Boesen, T., Pedersen, J.S., Spahn, C.M.*, et al.* (2006). Structure of eEF3 and the mechanism of transfer RNA release from the E-site. Nature *443*, 663-668.

Atkin, A.L., Altamura, N., Leeds, P., and Culbertson, M.R. (1995). The majority of yeast UPF1 co-localizes with polyribosomes in the cytoplasm. Molecular biology of the cell *6*, 611-625.

Blaha, G., Stanley, R.E., and Steitz, T.A. (2009). Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science *325*, 966-970.

Boel, G., Smith, P.C., Ning, W., Englander, M.T., Chen, B., Hashem, Y., Testa, A.J., Fischer, J.J., Wieden, H.J., Frank, J.*, et al.* (2014). The ABC-F protein EttA gates ribosome entry into the translation elongation cycle. Nature structural & molecular biology *21*, 143-151.

Brown, A., Shao, S., Murray, J., Hegde, R.S., and Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. Nature *524*, 493-496.

Budkevich, T., Giesebrecht, J., Altman, R.B., Munro, J.B., Mielke, T., Nierhaus, K.H., Blanchard, S.C., and Spahn, C.M. (2011). Structure and dynamics of the mammalian ribosomal pretranslocation complex. Molecular cell *44*, 214-224.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. Molecular cell *41*, 693-703.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nature structural & molecular biology *15*, 85-93.

Chen, B., Boel, G., Hashem, Y., Ning, W., Fei, J., Wang, C., Gonzalez, R.L., Jr., Hunt, J.F., and Frank, J. (2014). EttA regulates translation by binding the ribosomal E site and restricting ribosome-tRNA dynamics. Nature structural & molecular biology *21*, 152-159.

Cheng, Z., Muhlrad, D., Lim, M.K., Parker, R., and Song, H. (2007). Structural and functional insights into the human Upf1 helicase core. The EMBO journal *26*, 253-264.

Clerici, M., Mourao, A., Gutsche, I., Gehring, N.H., Hentze, M.W., Kulozik, A., Kadlec, J., Sattler, M., and Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. The EMBO journal *28*, 2293-2306.

Culbertson, M.R., Underbrink, K.M., and Fink, G.R. (1980). Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. Genetics *95*, 833-853.

Culver, G.M., Cate, J.H., Yusupova, G.Z., Yusupov, M.M., and Noller, H.F. (1999). Identification of an RNA-protein bridge spanning the ribosomal subunit interface. Science *285*, 2133-2136.

Czaplinski, K., Ruiz-Echevarria, M.J., Paushkin, S.V., Han, X., Weng, Y., Perlick, H.A., Dietz, H.C., Ter-Avanesyan, M.D., and Peltz, S.W. (1998). The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. Genes & development *12*, 1665-1677.

Czaplinski, K., Weng, Y., Hagan, K.W., and Peltz, S.W. (1995). Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation. RNA *1*, 610-623.

Dixon, W.J., Hayes, J.J., Levin, J.R., Weidner, M.F., Dombroski, B.A., and Tullius, T.D. (1991). Hydroxyl radical footprinting. Methods in enzymology *208*, 380-413.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85-88.

Dunkle, J.A., Wang, L., Feldman, M.B., Pulk, A., Chen, V.B., Kapral, G.J., Noeske, J., Richardson, J.S., Blanchard, S.C., and Cate, J.H. (2011). Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. Science *332*, 981-984.

Eyler, D.E., and Green, R. (2011). Distinct response of yeast ribosomes to a miscoding event during translation. RNA *17*, 925-932.

Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M.V., Yusupova, G., and Yusupov, M. (2014). Structural basis for the inhibition of the eukaryotic ribosome. Nature *513*, 517-522.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Molecular cell *51*, 35-45.

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., and Jacobson, A. (1993). Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. Proceedings of the National Academy of Sciences of the United States of America *90*, 7034-7038.

He, S.L., and Green, R. (2010). Visualization of codon-dependent conformational rearrangements during translation termination. Nature structural & molecular biology *17*, 465-470.

Heilek, G.M., and Noller, H.F. (1996). Site-directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S5. Science *272*, 1659-1662.

Holbrook, J.A., Neu-Yilik, G., Hentze, M.W., and Kulozik, A.E. (2004). Nonsense-mediated decay approaches the clinic. Nature genetics *36*, 801-808.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. The EMBO journal *27*, 736-747.

Jiang, H., Lei, R., Ding, S.W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC bioinformatics *15*, 182.

Joseph, S., Weiser, B., and Noller, H.F. (1997). Mapping the inside of the ribosome with an RNA helical ruler. Science *278*, 1093-1098.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nature reviews. Molecular cell biology *13*, 700-712.

Khatter, H., Myasnikov, A.G., Natchiar, S.K., and Klaholz, B.P. (2015). Structure of the human 80S ribosome. Nature *520*, 640-645.

Kielpinski, L.J., and Vinther, J. (2014). Massive parallel-sequencing-based hydroxyl radical probing of RNA accessibility. Nucleic acids research *42*, e70.

Leeds, P., Wood, J.M., Lee, B.S., and Culbertson, M.R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Molecular and cellular biology *12*, 2165-2177.

Lelivelt, M.J., and Culbertson, M.R. (1999). Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome. Molecular and cellular biology *19*, 6710-6719.

Losson, R., and Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. Proceedings of the National Academy of Sciences of the United States of America *76*, 5134-5137.

Marquardt, S., Hazelbaker, D.Z., and Buratowski, S. (2011). Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. Transcription *2*, 145-154.

Matheisl, S., Berninghausen, O., Becker, T., and Beckmann, R. (2015). Structure of a human translation termination complex. Nucleic acids research *43*, 8615-8626.

Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Human molecular genetics *10*, 99-105.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nature genetics *36*, 1073-1078.

Min, E.E., Roy, B., Amrani, N., He, F., and Jacobson, A. (2013). Yeast Upf1 CH domain interacts with Rps26 of the 40S ribosomal subunit. RNA *19*, 1105-1115.

Mitrovich, Q.M., and Anderson, P. (2000). Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in C. elegans. Genes & development *14*, 2173-2184.

Mohan, S., and Noller, H.F. (2017). Recurring RNA structural motifs underlie the mechanics of L1 stalk movement. Nature communications *8*, 14285.

Nierhaus, K.H. (1990). The allosteric three-site model for the ribosomal elongation cycle: features and future. Biochemistry *29*, 4997-5008.

Powers, T., and Noller, H.F. (1995). Hydroxyl radical footprinting of ribosomal proteins on 16S rRNA. RNA *1*, 194-209.

Preis, A., Heuer, A., Barrio-Garcia, C., Hauser, A., Eyler, D.E., Berninghausen, O., Green, R., Becker, T., and Beckmann, R. (2014). Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. Cell reports *8*, 59-65.

Rana, T.M., and Meares, C.F. (1991). Transfer of oxygen from an artificial protease to peptide carbon during proteolysis. Proceedings of the National Academy of Sciences of the United States of America *88*, 10578-10582.

Rehwinkel, J., Letunic, I., Raes, J., Bork, P., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. RNA *11*, 1530-1544.

Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2016). Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. Nucleic acids research *44*, 1944-1951.

Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., and Liu, J.O. (2010). Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. Nature chemical biology *6*, 209-217.

Schoenberg, D.R., and Maquat, L.E. (2012). Regulation of cytoplasmic mRNA decay. Nature reviews. Genetics *13*, 246-259.

Schuller, A.P., Wu, C.C., Dever, T.E., Buskirk, A.R., and Green, R. (2017). eIF5A Functions Globally in Translation Elongation and Termination. Molecular cell *66*, 194-205 e195.

Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., and Hegde, R.S. (2016). Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. Cell *167*, 1229-1240 e1215.

Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. Science *330*, 369-372.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proceedings of the National Academy of Sciences of the United States of America *108*, E1392-1398.

Shoemaker, C.J., and Green, R. (2012). Translation drives mRNA quality control. Nature structural & molecular biology *19*, 594-601.

Smith, J.E., Alvarez-Dominguez, J.R., Kline, N., Huynh, N.J., Geisler, S., Hu, W., Coller, J., and Baker, K.E. (2014). Translation of small open reading frames within unannotated RNA transcripts in Saccharomyces cerevisiae. Cell reports *7*, 1858-1866.

Takeshita, K., Forget, B.G., Scarpa, A., and Benz, E.J., Jr. (1984). Intranuclear defect in beta-globin mRNA accumulation due to a premature translation termination codon. Blood *64*, 13-22.

Triana-Alonso, F.J., Chakraburtty, K., and Nierhaus, K.H. (1995). The elongation factor 3 unique in higher fungi and essential for protein biosynthesis is an E site factor. The Journal of biological chemistry *270*, 20473-20478.

Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., and Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science *339*, 82-85.

Welch, E.M., and Jacobson, A. (1999). An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA. The EMBO journal *18*, 6134-6145.

Weng, Y., Czaplinski, K., and Peltz, S.W. (1998). ATP is a cofactor of the Upf1 protein that modulates its translation termination and RNA binding activities. RNA *4*, 205-214.

Wilson, K.S., and Noller, H.F. (1998). Mapping the position of translational elongation factor EF-G in the ribosome by directed hydroxyl radical probing. Cell *92*, 131-139.

Youngman, E.M., Brunelle, J.L., Kochaniak, A.B., and Green, R. (2004). The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. Cell *117*, 589-599.

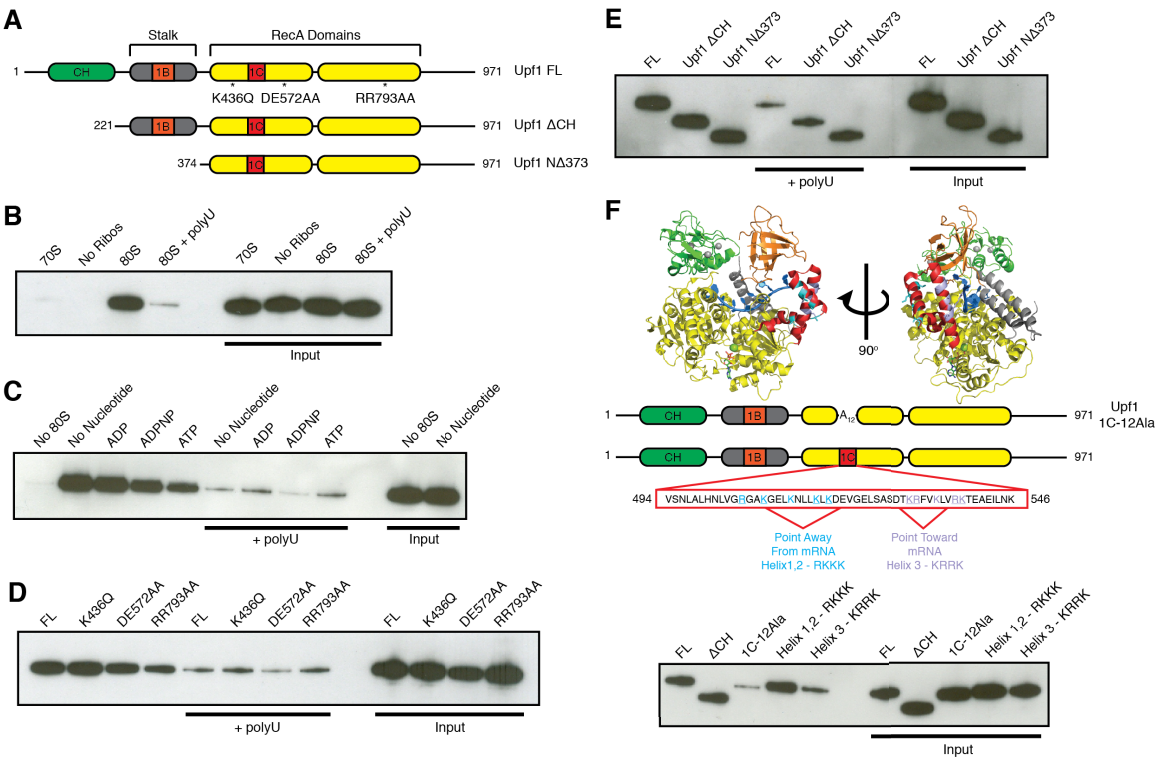

A
Stalk
RecA Domains
CH
1B
1C
K436Q DE572AA
RR793AA
Upf1 FL
Upf1 ΔCH
Upf1 NΔ373
B
70S
No Ribos
80S
80S + polyU
Input
C
No 80S
No Nucleotide
ADP
ADPNP
ATP
+ polyU
Input
D
FL
K436Q
DE572AA
RR793AA
+ polyU
Input
E
Upf1 ΔCH
Upf1 NΔ373
+ polyU
Input
F
90°
Upf1
1C-12Ala
494 VSNLALHNLVGRGAKGELKNLLKLKDEVGELSASDTKRFVKLVRKTEAEILNK 546
Point Away
From mRNA
Helix1,2 - RKKK
Point Toward
mRNA
Helix 3 - KRRK
ΔCH
1C-12Ala
Helix 1,2 - RKKK
Helix 3 - KRRK
Input

Figure 1

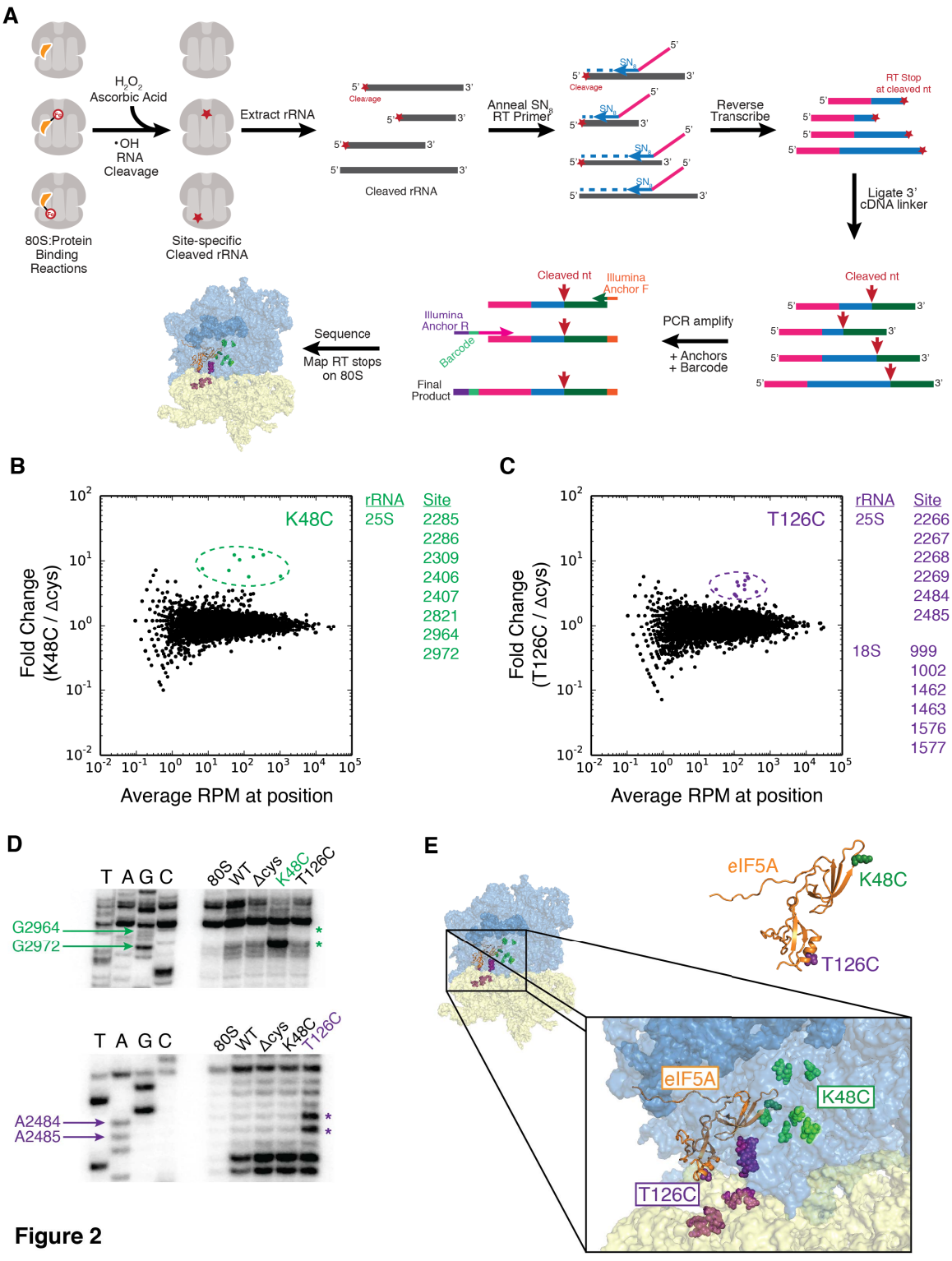

A
H2O2
Ascorbic Acid
•OH
RNA
Cleavage
80S:Protein
Binding
Reactions
Site-specific
Cleaved rRNA
Extract rRNA
Cleavage
Cleaved rRNA
Anneal SN8
RT Primer
Reverse
Transcribe
RT Stop
at cleaved nt
Ligate 3'
cDNA linker
PCR amplify
+ Anchors
+ Barcode
Cleaved nt
Illumina
Anchor F
Illumina
Anchor R
Barcode
Final
Product
Sequence
Map RT stops
on 80S
B
K48C
Fold Change
(K48C / Δcys)
Average RPM at position
rRNA
Site
25S
2285
2286
2309
2406
2407
2821
2964
2972
C
T126C
Fold Change
(T126C / Δcys)
Average RPM at position
rRNA
Site
25S
2266
2267
2268
2269
2484
2485
18S
999
1002
1462
1463
1576
1577
D
T A G C
80S
WT
Δcys
K48C
T126C
G2964
G2972
A2484
A2485
E
eIF5A
K48C
T126C

Figure 2

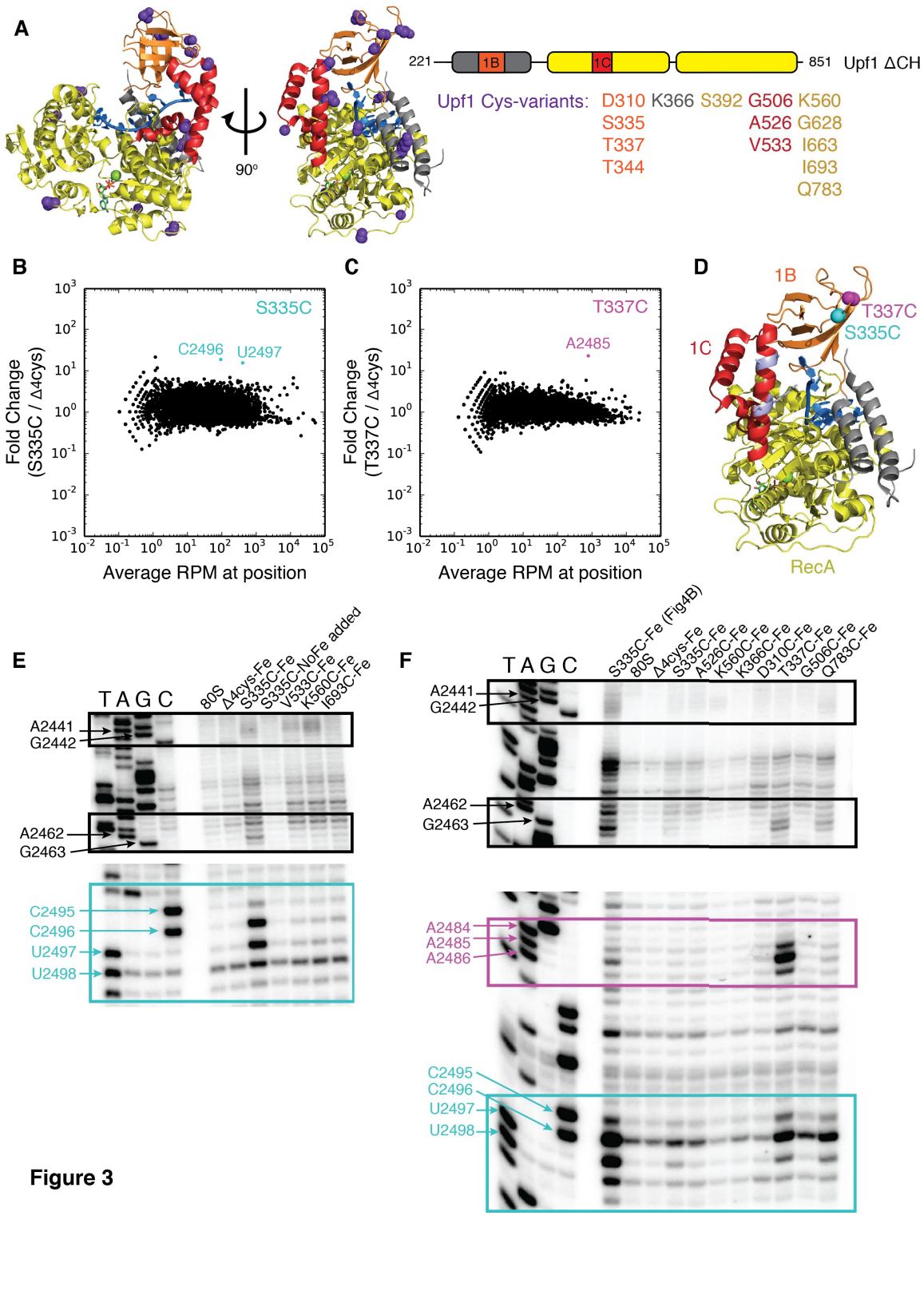

A
90°
221
1B
1C
851
Upf1 ΔCH
Upf1 Cys-variants: D310 K366 S392 G506 K560
S335 A526 G628
T337 V533 I663
T344 I693
Q783
B
S335C
C2496
U2497
Fold Change (S335C / Δ4cys)
Average RPM at position
C
T337C
A2485
Fold Change (T337C / Δ4cys)
Average RPM at position
D
1B
T337C
S335C
1C
RecA
E
T A G C
80S
Δ4cys-Fe
S335C-Fe
S335C-NoFe added
V533C-Fe
K560C-Fe
I693C-Fe
A2441
G2442
A2462
G2463
C2495
C2496
U2497
U2498
F
T A G C
S335C-Fe (Fig4B)
80S
Δ4cys-Fe
S335C-Fe
A526C-Fe
K560C-Fe
K366C-Fe
D310C-Fe
T337C-Fe
G506C-Fe
Q783C-Fe
A2441
G2442
A2462
G2463
A2484
A2485
A2486
C2495
C2496
U2497
U2498

Figure 3

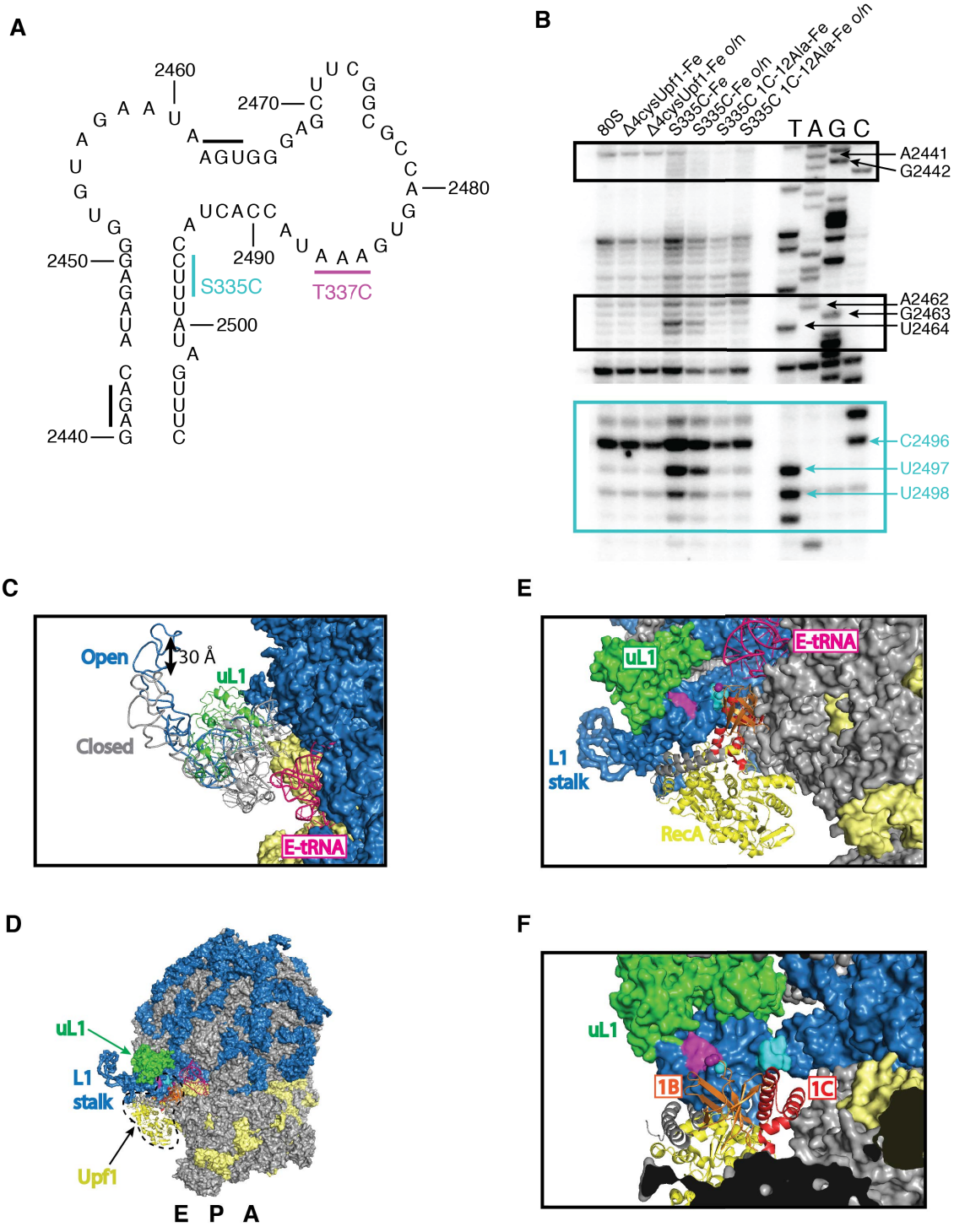

A
2460
2470
2480
2490
2500
2450
2440
S335C
T337C
B
80S
Δ4cysUpf1-Fe
Δ4cysUpf1-Fe o/n
S335C-Fe
S335C-Fe o/n
S335C 1C-12Ala-Fe
S335C 1C-12Ala-Fe o/n
T A G C
A2441
G2442
A2462
G2463
U2464
C2496
U2497
U2498
C
Open
30 Å
uL1
Closed
E-tRNA
E
uL1
E-tRNA
L1
stalk
RecA
D
uL1
L1
stalk
Upf1
E P A
F
uL1
1B
1C

Figure 4

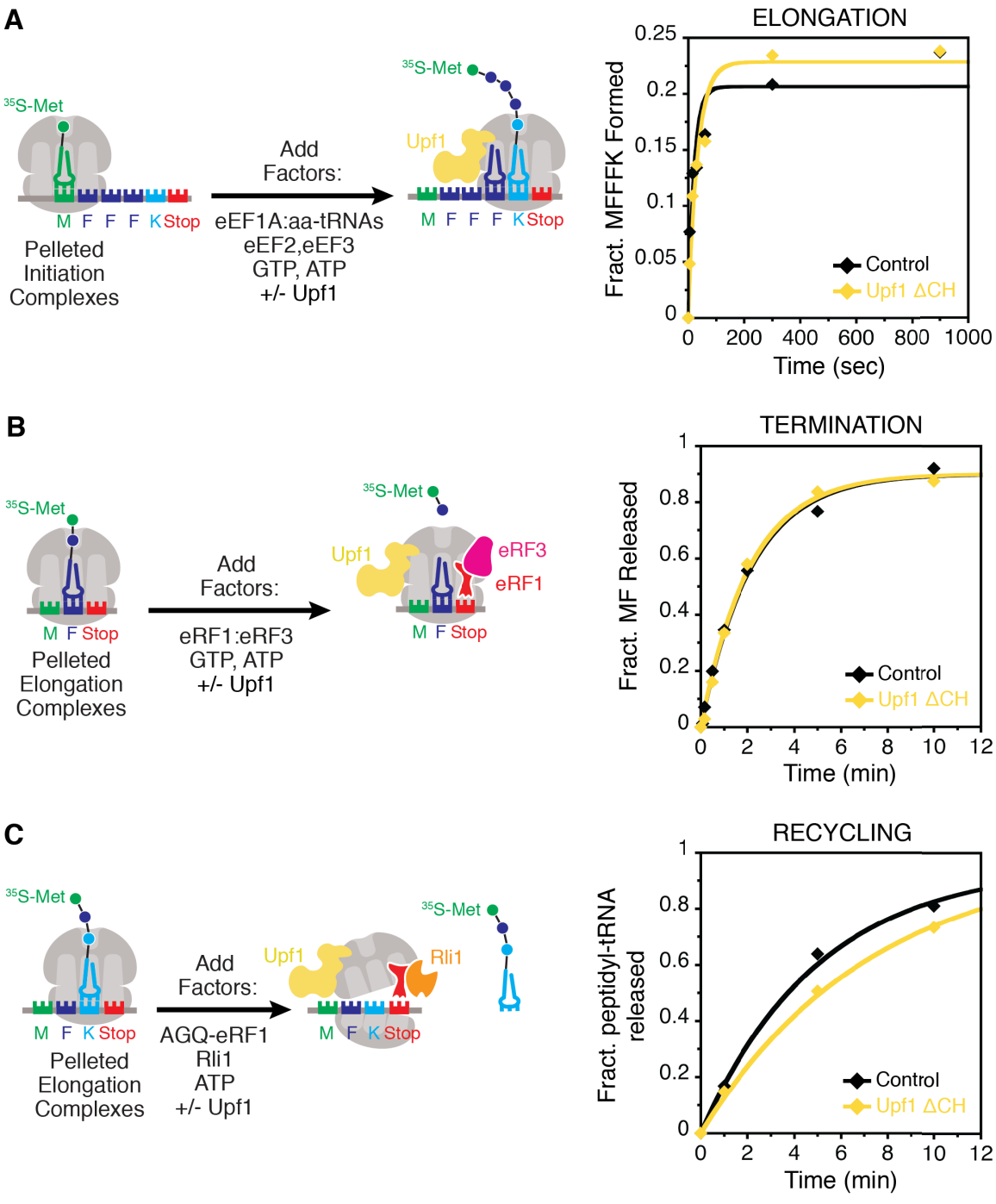
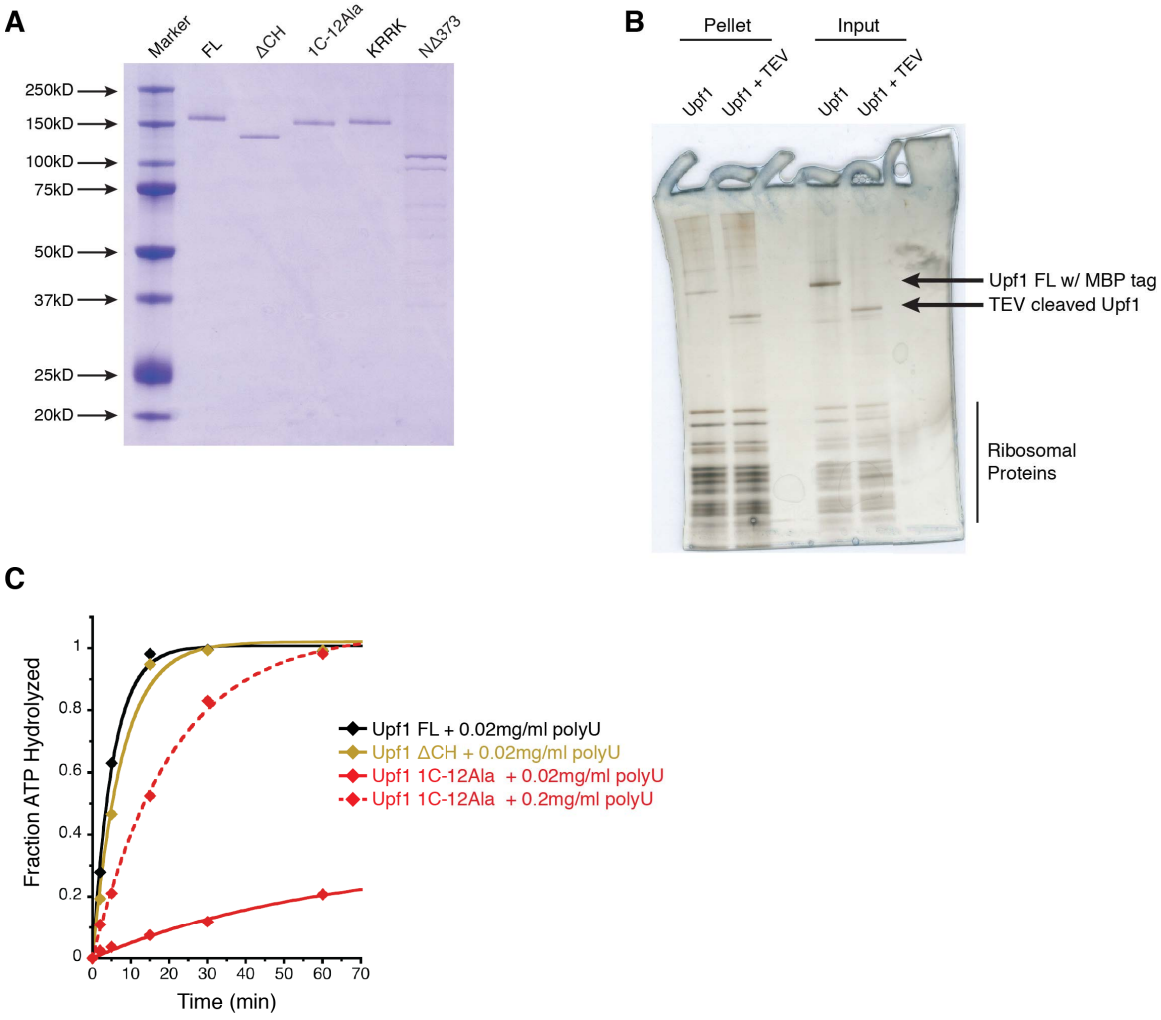
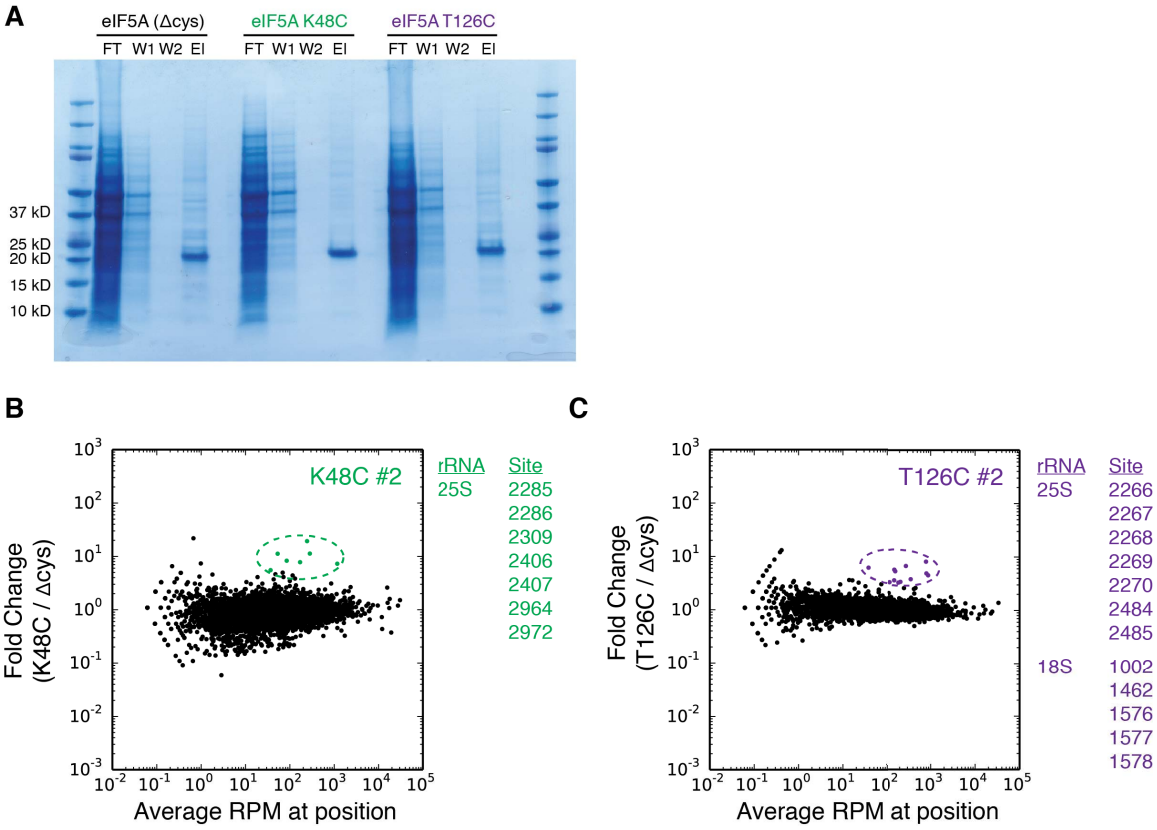

A
35S-Met
M F F F K Stop
Pelleted
Initiation
Complexes
Add
Factors:
eEF1A:aa-tRNAs
eEF2,eEF3
GTP, ATP
+/- Upf1
35S-Met
Upf1
M F F F K Stop
ELONGATION
Fract. MFFFK Formed
0.25
0.2
0.15
0.1
0.05
0
0 200 400 600 800 1000
Time (sec)
Control
Upf1 ΔCH
B
35S-Met
M F Stop
Pelleted
Elongation
Complexes
Add
Factors:
eRF1:eRF3
GTP, ATP
+/- Upf1
35S-Met
Upf1
eRF3
eRF1
M F Stop
TERMINATION
Fract. MF Released
1
0.8
0.6
0.4
0.2
0
0 2 4 6 8 10 12
Time (min)
Control
Upf1 ΔCH
C
35S-Met
M F K Stop
Pelleted
Elongation
Complexes
Add
Factors:
AGQ-eRF1
Rli1
ATP
+/- Upf1
Upf1
Rli1
35S-Met
M F K Stop
RECYCLING
Fract. peptidyl-tRNA
released
1
0.8
0.6
0.4
0.2
0
0 2 4 6 8 10 12
Time (min)
Control
Upf1 ΔCH

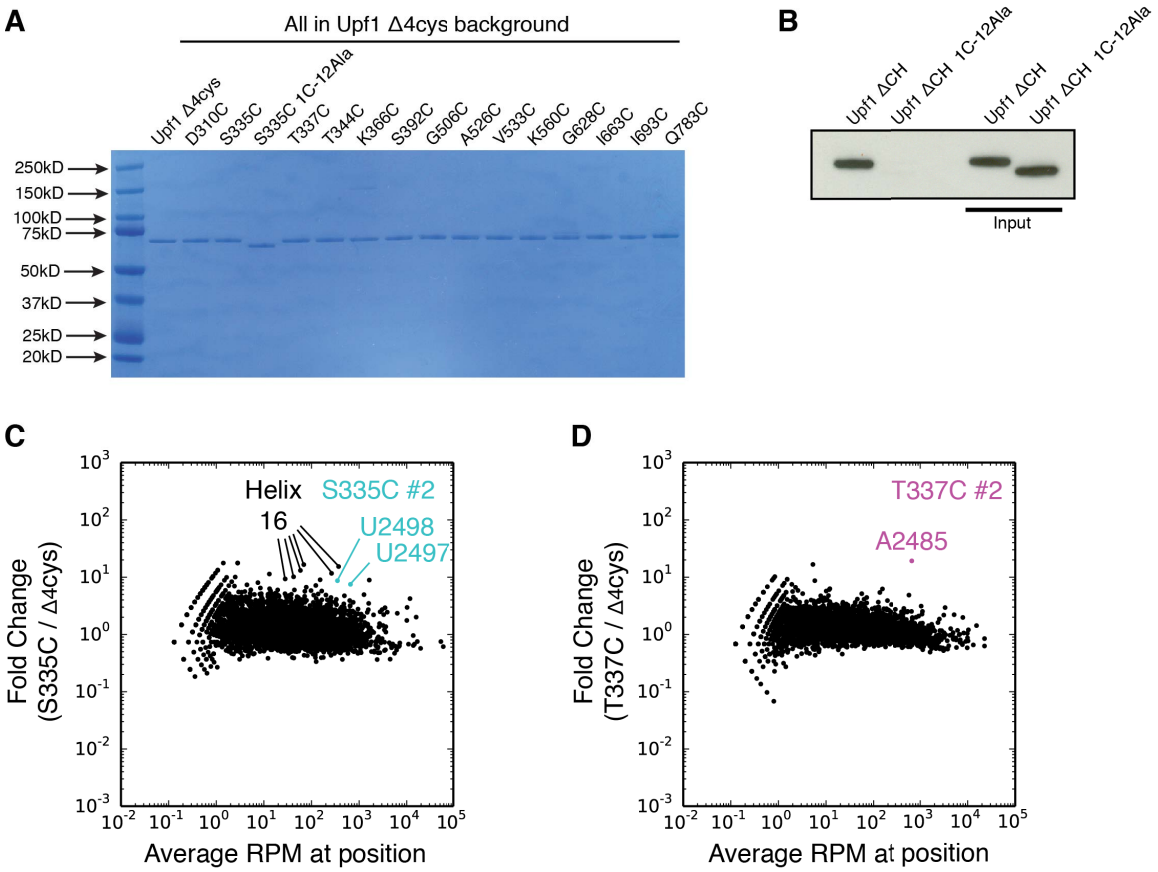
Figure 5



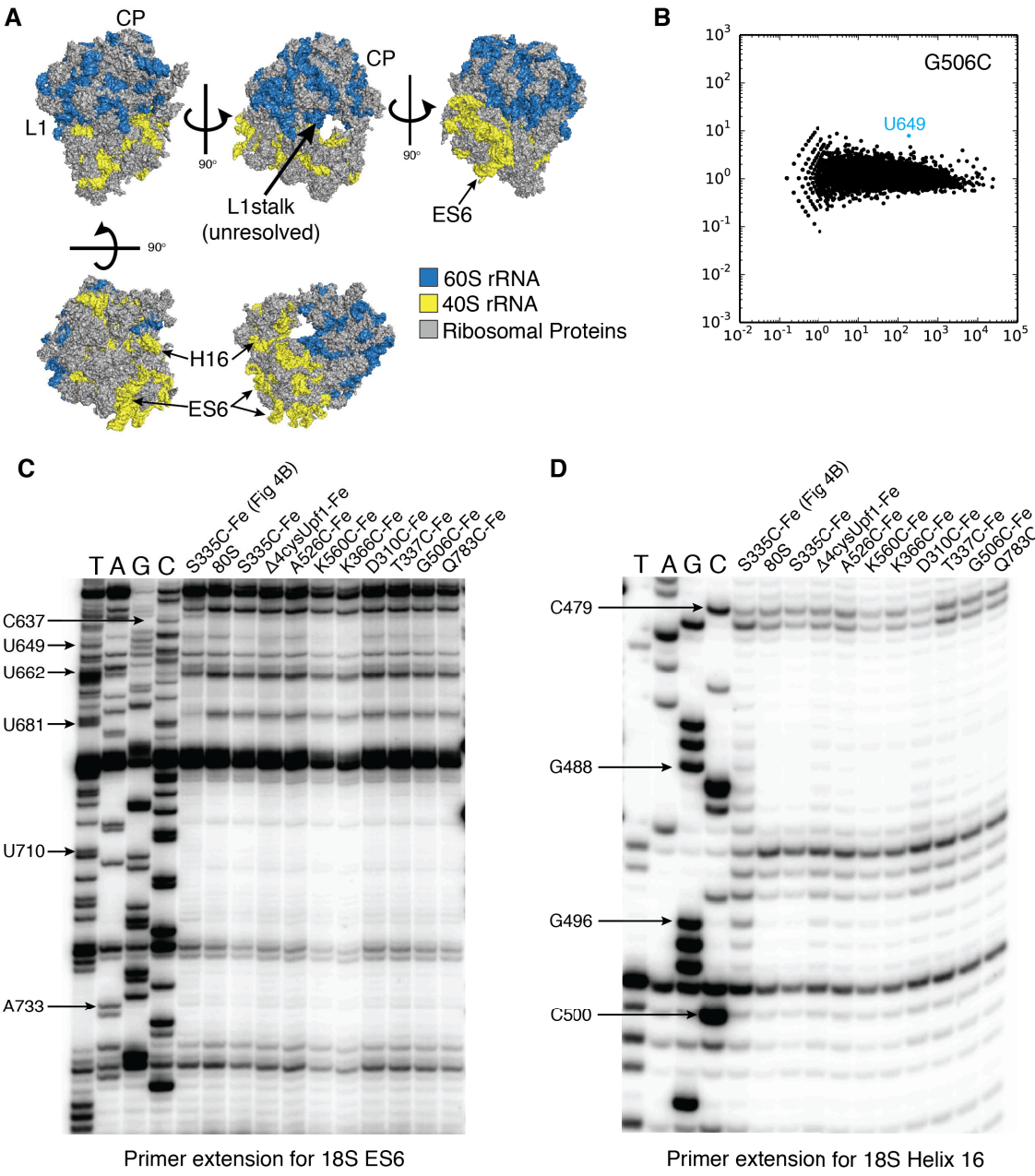
Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3



Primer extension for 18S ES6

Primer extension for 18S Helix 16

Supplemental Figure 4

Table S1

Oligo Name	Use	Sequence (5' --> 3')	Illumina Index Rev. Comp.
oAS510	18S Helix 16 probing	GTATTTACATTGTACTCATTCC	N/A
oAS511	18S ES6 probing	GCTAATATATTCGAGCAATACG	N/A
oAS517	25S Helix 93 probing	CTGACTTAGAGGCGTTCAGCC	N/A
oAS519	25S L1 stalk probing	AGCTCCGCTTCATTGAATAAG	N/A
oAS544	HTS-BABE RT	AGACGTGTGCTCTTCCGATCTNNNNNNNNS	N/A
oAS545	HTS-BABE Circligase	pNNNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGC-3′NHC3	N/A
oBZ287	Forward Illumina Oligo	AATGATACGGCGACCACCGAGATCTACAC	N/A
oASBar 01	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATaaacctGTGACTGGAGTTCAGACGTGTGCTCTTCCG	AGGTTT
oASBar 02	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATctcaggGTGACTGGAGTTCAGACGTGTGCTCTTCCG	CCTGAG
oASBar 03	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATgggttcGTGACTGGAGTTCAGACGTGTGCTCTTCCG	GAACCC
oASBar 04	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATctctgaaGTGACTGGAGTTCAGACGTGTGCTCTTCCG	TTCAGA
oASBar 05	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATacagcaGTGACTGGAGTTCAGACGTGTGCTCTTCCG	TGCTGT
oASBar 06	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATcaccgtGTGACTGGAGTTCAGACGTGTGCTCTTCCG	ACGGTG
oASBar 07	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATgtgatgGTGACTGGAGTTCAGACGTGTGCTCTTCCG	CATCAC
oASBar 08	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATtgttacGTGACTGGAGTTCAGACGTGTGCTCTTCCG	GTAACA
oASBar 09	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATagatccGTGACTGGAGTTCAGACGTGTGCTCTTCCG	GGATCT
oASBar 10	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAGATcccggaGTGACTGGAGTTCAGACGTGTGCTCTTCCG	TCCGGG

oASBar 11	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATgagcttGTGACTGGAGTTCAGACG TGTGCTCTTCCG	AAGCTC
oASBar 12	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATtttaagGTGACTGGAGTTCAGACG TGTGCTCTTCCG	CTTAAA
oASBar 13	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATaacttgGTGACTGGAGTTCAGACG TGTGCTCTTCCG	CAAGTT
oASBar 14	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATctggacGTGACTGGAGTTCAGAC GTGTGCTCTTCCG	GTCCAG
oASBar 15	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATggtccaGTGACTGGAGTTCAGAC GTGTGCTCTTCCG	TGGACC
oASBar 16	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATtcaagtGTGACTGGAGTTCAGACG TGTGCTCTTCCG	ACTTGA
oASBar 17	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATaccattGTGACTGGAGTTCAGACG TGTGCTCTTCCG	AATGGT
oASBar 18	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATcagtagGTGACTGGAGTTCAGAC GTGTGCTCTTCCG	CTACTG
oASBar 19	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATgttgccGTGACTGGAGTTCAGACG TGTGCTCTTCCG	GGCAAC
oASBar 20	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATtgacgaGTGACTGGAGTTCAGAC GTGTGCTCTTCCG	TCGTCA
oASBar 21	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATagcctaGTGACTGGAGTTCAGAC GTGTGCTCTTCCG	TAGGCT
oASBar 22	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATccgaatGTGACTGGAGTTCAGAC GTGTGCTCTTCCG	ATTCGG
oASBar 23	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATgattcgGTGACTGGAGTTCAGACG TGTGCTCTTCCG	CGAATC
oASBar 24	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATttaggcGTGACTGGAGTTCAGACG TGTGCTCTTCCG	GCCTAA
oASBar 25	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATatgctgGTGACTGGAGTTCAGACG TGTGCTCTTCCG	CAGCAT
oASBar 26	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATacgtcgGTGACTGGAGTTCAGAC GTGTGCTCTTCCG	CGACGT
oASBar 27	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATagctgcGTGACTGGAGTTCAGAC	GCAGCT

		GTGTGCTCTTCCG	
oASBar 28	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATatcgtaGTGACTGGAGTTCAGACG TGTGCTCTTCCG	TACGAT
oASBar 29	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATtggtcaGTGACTGGAGTTCAGACG TGTGCTCTTCCG	TGACCA
oASBar 30	Illumina Reverse Index	CAAGCAGAAGACGGCATACGAG ATattggcGTGACTGGAGTTCAGACG TGTGCTCTTCCG	GCCAAT

Chapter IV: Conclusion

In this work I have detailed my investigations of the function of two E-site binding proteins in translation. Nonetheless, many outstanding questions remain. Some of the more pressing issues are described below.

Role of eIF5A hypusine modification

In eukaryotes, eIF5A is an abundant and essential protein that we know to function in the stimulation of global elongation and termination (Schuller et al., 2017). In the course of my studies, I observed that eIF5A stimulates the translation of many ribosome-stalling motifs, but that its hypusine post-translational modification is only critical for the translation of poly-proline stretches. Recent structural data have suggested that the hypusine moiety makes interactions with the CCA-end of the peptidyl-tRNA, stabilizing it for nucleophilic attack at the ribosome peptidyl-transferase center (Schmidt et al., 2016).

Why might the hypusine modification be necessary for progression through poly-proline and not other stalling motifs, given the interaction is made through a conserved tRNA element? One possibility is that poly-proline stalls are caused not only by slow reaction kinetics, but also by an inhibitory conformation that poly-Pro adopts in the ribosome exit tunnel. In this model, interactions between the hypusine moiety and the tRNA would create a more favorable peptide conformation, while the body of eIF5A would be responsible for promoting fast peptidyl-transfer kinetics. A deeper investigation into the structure of eIF5A-bound ribosomes stalled at poly-proline and other stalling motifs, in addition to identification of the *in vivo* targets of unmodified eIF5A (potentially

by ribosome profiling), will be required to begin to tease out the precise function of eIF5A-hypusine modification.

Function of Upf1:80S interaction in NMD

In eukaryotes, the nonsense-mediated decay (NMD) pathway selectively degrades mRNAs containing a premature termination codon (PTC) (Kervestin and Jacobson, 2012). This quality control mechanism is dependent on the three Upf proteins: Upf1, Upf2, and Upf3. Upf1 is the primary, catalytic component of NMD and a member of the SF1 helicase family. In contrast, Upf2 and Upf3 are thought to act primarily as binding partners that regulate Upf1 location or function. In this work, I have discovered that the ribosome-binding site for Upf1 is the L1 stalk near the ribosomal E site. This region of the ribosome has been of great interest to the ribosome field, as it has been shown to adopt several unique conformations that correlate with the translation status of the ribosome (Mohan and Noller, 2017). While my observations are supported by other studies that have suggested an interaction for Upf1 near the ribosomal E site (Min et al., 2013), this binding site remains somewhat puzzling.

A key requirement for NMD is that an mRNA contains a PTC. From the ribosome's perspective, this recognition occurs in the A site, where the release factor eRF1 binds (Brown et al., 2015). It is possible, however, that Upf1 affects communication between the ribosome E and A sites by some long-range interaction, as has been shown for other protein factors, such as eEF3 (Andersen et al., 2006; Triana-Alonso et al., 1995). Alternatively, Upf1 may serve as a binding platform at the L1 stalk for other factors required to initiate NMD. Both *in vivo* and *in vitro* approaches will be

required to understand the functional consequence of this ribosomal interaction. A deeper understanding of the Upf1:80S interaction could be achieved by cryoEM in the near future, if this complex can be stably isolated (e.g., addition of ADP or use of an ATPase-inactive mutant) and the L1 stalk can be accurately resolved. Experiments using ribosome profiling in NMD-null backgrounds, or reporters to probe both NMD and translation termination efficiency will be required to understand how Upf1 can sense PTCs and what other factors are involved in this recognition.

A more fundamental question, however, may simply be to ask what factors are involved in translation termination? At the protein level, recent work with the poly(A)-binding protein (PABP) has implicated the distance between PABP and a termination codon as a determinant of termination efficiency (Swart et al., 2016). Biochemical data in a reconstituted translation system supports this connection and suggests that PABP could alter the conformation of a terminating ribosome bound to eRF1:eRF3 (Ivanov et al., 2016). Understanding the details of this interaction and its effect on termination efficiency will be an important step moving forward. Also, since the atlas of mRNA binding proteins is quite large (Castello et al., 2012), it is possible that other factors could similarly affect termination. Coordination (or lack thereof) between these factors, the terminating ribosome, and Upf1 could be the key to understanding NMD.

REFERENCES

Andersen, C.B., Becker, T., Blau, M., Anand, M., Halic, M., Balar, B., Mielke, T., Boesen, T., Pedersen, J.S., Spahn, C.M.*, et al.* (2006). Structure of eEF3 and the mechanism of transfer RNA release from the E-site. Nature *443*, 663-668.

Brown, A., Shao, S., Murray, J., Hegde, R.S., and Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. Nature *524*, 493-496.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M.*, et al.* (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell *149*, 1393-1406.

Ivanov, A., Mikhailova, T., Eliseev, B., Yeramala, L., Sokolova, E., Susorov, D., Shuvalov, A., Schaffitzel, C., and Alkalaeva, E. (2016). PABP enhances release factor recruitment and stop codon recognition during translation termination. Nucleic acids research *44*, 7766-7776.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nature reviews Molecular cell biology *13*, 700-712.

Min, E.E., Roy, B., Amrani, N., He, F., and Jacobson, A. (2013). Yeast Upf1 CH domain interacts with Rps26 of the 40S ribosomal subunit. Rna *19*, 1105-1115.

Mohan, S., and Noller, H.F. (2017). Recurring RNA structural motifs underlie the mechanics of L1 stalk movement. Nature communications *8*, 14285.

Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2016). Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. Nucleic acids research *44*, 1944-1951.

Schuller, A.P., Wu, C.C., Dever, T.E., Buskirk, A.R., and Green, R. (2017). eIF5A Functions Globally in Translation Elongation and Termination. Molecular cell *66*, 194-205 e195.

Swart, E.C., Serra, V., Petroni, G., and Nowacki, M. (2016). Genetic Codes with No Dedicated Stop Codon: Context-Dependent Translation Termination. Cell *166*, 691-702.

Triana-Alonso, F.J., Chakraburtty, K., and Nierhaus, K.H. (1995). The elongation factor 3 unique in higher fungi and essential for protein biosynthesis is an E site factor. The Journal of biological chemistry *270*, 20473-20478.

Appendix I: eIF5A is a common contaminant of yeast protein purifications and a potent stimulator of translation termination

Abstract

During the course of my work with Upf1, an essential factor in the nonsense-mediated decay pathway, I discovered eIF5A as a potent stimulator of translation termination. My investigations began by using an *in vitro* reconstituted translation system to determine the effect of Upf1 on translation termination. Initial results suggested that Upf1 stimulated translation termination, and specifically the catalytic function of eRF1 to promote peptidyl-hydrolysis. After purification of several mutant and truncated Upf1 constructs, I determined that the stimulatory activity was instead due to a contaminating factor specific to protein purifications from *S. cerevisiae* cells. Using a combination of biochemical fractionation, activity assays in a reconstituted translation system, and mass spectrometry, I determined this contaminating activity was due to eIF5A. Subsequent analysis of other proteins purified from *S. cerevisiae* (such as eEF2) revealed that eIF5A is a common contaminant of yeast protein purifications. Special precautions must be taken in order to purify yeast proteins in the absence of eIF5A to study both translation elongation and termination *in vitro*.

Parts of this appendix were published in Molecular Cell 66, 194-205 (2017)

RESULTS

Upf1 stimulates translation termination

Using an *in vitro* reconstituted translation system, we investigated the role of Upf1 in translation termination. Programmed termination complexes composed of 80S yeast ribosomes on an mRNA with [^{35}S]-Met-Phe-tRNA^{Phe} in the P (peptidyl) site and a stop codon (UAA) in the A site were used to follow translation termination over time. Termination complexes were incubated with release factors eRF1:eRF3 and the rate of hydrolysis of [^{35}S]-Met-Phe (MF) peptide from the peptidyl-tRNA was followed by electrophoretic TLC (Figure 1A). The rate constant (k_{obs}) for eRF1:eRF3 stimulated peptide release was 0.6 min^{-1}, consistent with previous reports (Eyler and Green, 2011; Shoemaker and Green, 2011). When Upf1 was added to the termination reaction, the rate of peptidyl-hydrolysis by eRF1:eRF3 was stimulated 14-fold to a rate of 8.4 min^{-1} (Figure 1B). Addition of 3-fold more Upf1 resulted in no further stimulation indicating that the reaction was saturated and, at a minimum, that there are substantial effects on k_{cat} (data not shown). Similar stimulation of termination activity by Upf1 was observed in a reaction with only eRF1 (in the absence of eRF3); the rate of release by Upf1:eRF1 was 6-fold higher than the rate with eRF1 alone. Additionally, the non-hydrolyzable GTP analog (GDPNP) strongly inhibited the termination reaction with eRF1:eRF3:Upf1, consistent with the idea that the release of eRF3 from the termination complex by GTP hydrolysis is required prior to the peptide hydrolysis reaction (Eyler et al., 2013; Shoemaker and Green, 2011).

In light of the known coupling between termination and recycling in eukaryotes (Shoemaker and Green, 2011) and because of proposed roles for Upf1 in modulating

ribosome recycling (Amrani et al., 2006; Celik et al., 2015), we next evaluated the ability of Upf1 to promote subunit dissociation. In these experiments, we monitored the rate of subunit dissociation of termination complexes incubated with canonical termination and recycling factors (eRF1:eRF3 and Rli1), or with combinations of these factors and Upf1. The previously characterized, catalytically inactive, eRF1 mutant (GGQ→AGQ) was used to more simply follow the recycling reaction (it prevents hydrolysis of the MF peptide by the termination factor but still results in subunit dissociation). As previously reported, Rli1 (ABCE1 in higher eukaryotes) substantially increases the natural rate of ribosome recycling in our assay while Upf1 has no discernible effect relative to background (AGQ:eRF3) (Figure 1C).

We next asked whether there was a mechanistic connection between Upf1 helicase activity or ribosome binding, and this stimulatory function in termination. This possibility was tested by asking whether any of Upf1 variants including those defective in ATP binding (K436Q), ATP hydrolysis (DE572AA), RNA binding (RR793AA), or those lacking one or several domains could enhance the rate of the peptide release reaction (Figure 1D); we note that only the 1C deletion failed to bind ribosomes in our pelleting assay (data not shown). To our surprise, we found that all of these constructs were able to fully promote peptidyl-hydrolysis under saturating (k_{cat}) conditions identified for full-length Upf1 (Figure 1E). Together, these results fail to establish any direct connection between ribosome binding or helicase function by Upf1 and its contribution to peptide release.

Identification of contaminating activity in *S. cerevisiae* protein purifications

In light of this unexplained mechanism for Upf1 stimulation, we decided to purify several truncations of Upf1 lacking the helicase core (such as the Upf1 CH domain only), as well as other RNA helicases (Dhh1 for example). Performing the same *in vitro* termination assays with both of these constructs revealed that stimulation of translation termination was not due to Upf1, but rather some contaminant of these purifications (data not shown). To better investigate the cause of this contaminating activity, we purified a His-MBP construct from both *S. cerevisiae* and *E. coli* cells as all Upf1 constructs contained these tags for purification. Upon addition of these His-MBP proteins we found that MBP purified from *S. cerevisiae* robustly stimulated translation termination while MBP purified from *E. coli* did not (Figure 2).

To further identify the source of this activity, we tested whether the activity could be attributed to a protein, RNA, or small molecule contaminant using the *in vitro* termination assay. We tested a multitude of conditions including: 1) various buffer components to identify if a small molecule was the cause of such activity, 2) whether the contaminant was an RNA species by adding RNAse (data not shown), or 3) whether the contaminant was proteinacious by both heat treament and addition of proteinase K (Figure 3). From these data we determined that the contaminating activity was due to a proteinacious molecule that was heat-insensitive. As heat-sensitivity was tested by boiling the protein fraction at 95 degrees for 15 minutes and then cooling to room temperature, it is also possible the protein was able to re-fold after heat-treatment. Taken together, this data suggested that a protein contaminant specific to the yeast protein

purifications (by Ni-NTA and amylose resin chromatography) was responsible for the stimulatory activity.

eIF5A is a potent stimulator of translation termination

To identify the protein contaminant, we tried to enrich for activity by isolating fractions from affinity (Ni-NTA and amylose) and ion exchange chromatography. This course of action was not productive, as fractions from all these columns were able to stimulate the *in vitro* termination reaction (data not shown). While some fractions contained more activity than others, it was not clear how to compare these fractions to identify the protein of interest. Instead, we performed ESI mass spectrometry to identify all protein components of several purifications that contained termination stimulatory activity (data not shown). After analysis of the peptides identified from a tryptic digestion, we identified eIF5A as a candidate for this contaminating activity.

eIF5A is a small and highly-abundant protein in yeast cells (Kulak et al., 2014) that was shown to stimulate peptide bond formation at Pro-Pro sequences (Gutierrez et al., 2013), interacting with the peptidyl-tRNA through the ribosomal E site (Schmidt et al., 2016). Because of the similarities between peptide-bond formation and termination (both nucleophilic attack by an A-site molecule at the peptidyl-transferase center of the ribosome), we identified eIF5A as a candidate for this stimulation of translation termination. To investigate this directly we purified recombinant eIF5A from *E. coli* and tested its function in the *in vitro* termination assay (Figure 4). To our surprise, eIF5A showed robust stimulation of eRF1-mediated translation termination, and was also heat-insensitive. Additionally, we could attribute the stimulatory activity to a unique post-

translational modification on eIF5A called hypusine (Park et al., 1981). Taken together these data suggested that the contaminating activity in our yeast purifications was due to hypusinated eIF5A.

eIF5A is contaminant of other *S. cerevisiae* protein purifications

Because eIF5A is such an abundant protein in cells and co-purified with the Upf1 constructs after both nickel and amylose chromatography, we were worried that eIF5A could also be a contaminant of other factors used in our *in vitro* reconstituted system. To test this we performed *in vitro* elongation reactions to study the rate of Pro-Pro bond formation, a kinetically slow reaction that was shown to require eIF5A. To our surprise we found that we could robustly make an MPPK peptide even without adding additional eIF5A, suggestive of eIF5A contamination in our preparations of eEF2 and eEF3 (data not shown). To solve this issue, we purified these factors with a final gel filtration chromatography step, and isolated fractions individually (rather than pooling the fractions for concentration). Individual fractions were concentrated and tested for MPPK formation to identify fractions that did (or did not) contain eIF5A (Figure 5). We found that early fractions of eEF2 from an S200 column had robust eEF2 function without contaminating eIF5A, while later fractions were in fact contaminated. This analysis is the only way to identify eIF5A contamination as Western blotting did not provide a signal for eIF5A, even in these early fractions (data not shown). Taken together these results suggest that eIF5A is a robust and common contaminant of proteins isolated from *S. cerevisiae* cells.

DISCUSSION

Through a combination of *in vitro* biochemistry and mass spectrometry, we identified eIF5A as a robust contaminant of protein purifications from *S. cerevisiae*. eIF5A is abundant, heat-insensitive, and co-purifies by many chromatographic methods including nickel, amylose, and ion exchange. In light of these issues, special considerations must be taken to avoid eIF5A contamination when necessary. To isolate proteins in the absence of eIF5A, we found careful fractionation and analysis of Pro-Pro bond formation as the most robust method to identify contamination. Moving forward, it might be useful to include an affinity tag (such as FLAG or HA) at the genomic locus of eIF5A in *S. cerevisiae* strains used for protein purification. In this way, the eIF5A could be "purified out" using an affinity resin, to ensure the purity of the sample.

MATERIALS AND METHODS

Upf1 Expression and Purification

The Upf1 gene from S. cerevisiae (NAM7) was cloned into pYES-DEST52 (Invitrogen) with TEV protease-cleavable N-terminal His_7 and C-terminal maltose binding protein (MBP) tags for affinity purification. All variants of Upf1 (point mutations and truncations) were prepared by sub-cloning from this parent plasmid using Topo Directional cloning (Invitrogen) or QuickChange mutagenesis (Stratagene). The pYES-DEST52 plasmid contains a *GAL1* inducible promoter for expression of Upf1 in *S. cerevisiae*. All Upf1 proteins were expressed in *S. cerevisiae* JC287 (MATa, *ade2-1, his3, leu2, trp1, ura3, pep4::HIS3, prb1::HIS3, prc1::HIS3*) (provided by Dr. Jeff Coller) with 2% galactose. Cell pellets were harvested, flash frozen in buffer containing 25 mM Tris, 500 mM NaCl, 10% glycerol, 1 mM MgCl2, 1 mM ZnCl2, and 5 mM BME (A500), and lysed in a liquid nitrogen Freezer/Mill (SPEX Sample Prep, LLC). Cell lysates were purified over an Amylose column (NEB) followed by a HisTrapFF column (GE Healthcare) to yield pure full-length Upf1 protein. Protein fractions were pooled, concentrated, exchanged into buffer A150 (same as A500 but with 150 mM NaCl) and stored at -80^{o}C.

Purification of translation factors

Translation initiation factors eIF2, eIF1, eIF1A, eIF5, and eIF5b were purified from both *S. cerevisiae* (eIF2) and *E. coli* using reported procedures. Purification of eEF1A from *S. cerevisiae* also followed a previously reported protocol.

During purification of eEF2 and eEF3 from *S. cerevisiae* we identified the presence of low-level (though highly active) contaminating eIF5A. Endogenous eIF5A is very abundant, and can be enriched from yeast lysate by both Ni-NTA and amylose affinity chromatography. In order to isolate eEF2 and eEF3 in the absence of eIF5A, we took great care to sacrifice any contaminated fractions during purification to yield a highly pure protein. After purification, individual fractions were tested for MPPK synthesis in the absence of eIF5A as a test for contamination. Those fractions that did not stimulate MPPK formation were used in subsequent assays.

eEF2 was purified using a C-terminal His_6 tag from strain TKY675. eEF2 was purified by Ni-NTA chromatography in buffer containing 20 mM HEPES pH 7.4, 10% glycerol, 300 mM KCl, and 5 mM 2-mercaptoethanol. Then the protein was diluted to 30 mM KCl and further purified by cation exchange chromatography using a gradient from 30-150 mM KCl. Fractions containing eEF2 (and no eIF5A) were pooled, concentrated, and further purified by gel filtration using a HiPrep 26/60 Sephacryl S-200 column (GE Healthcare Life Sciences) for best separation of eEF2 and contaminant eIF5A. Fractions containing eEF2 were individually concentrated, and stored in gel filtration buffer containing 20 mM HEPES pH 7.4, 100 mM KOAc pH 7.6, 10% glycerol, and 2 mM DTT.

eEF3 was purified using a N-terminal His_6 tag from strain TKY702. eEF3 was purified by Ni-NTA chromatography in buffer containing 25 mM Tris pH 7.5, 10 mM Imidazole, 1 M KCl, 10% glycerol, and 5 mM 2-mercaptoethanol. After elution, protein was further purified by gel filtration using a Superose 6 10/300GL column (GE Healthcare Life Sciences). Fractions containing eEF3 were individually concentrated, and

stored in gel filtration buffer containing 20 mM HEPES pH 7.4, 100 mM KOAc pH 7.6, 10% glycerol, and 2 mM DTT.

Eukaryotic release factor eRF1 (and AGQ mutant) was purified from *E. coli* with a His_6 tag as previously described. Release factor eRF3 was purified from *E. coli* using the IMPACT Protein Purification System (New England Biolabs) as previously described.

Expression and purification of recombinant eIF5A from *E. coli* was performed as previously reported. Briefly, a plasmid expressing His_6-eIF5A, or a plasmid co-expressing His_6-eIF5A with modification enzymes Dys1 and Lia1 was transformed into BL21-CodonPlus (DE3)-RIPL cells (Agilent). Proteins were expressed and purified by Ni-NTA affinity chromatography and anion exchange chromatography.

In vitro Termination and Subunit Separation

Ribosome complexes encoding Met-Phe-$tRNA^{Phe}$ in the P site and a stop codon (UAA) in the A site were formed and purified as previously described. Termination assays were performed in 1X Buffer E (20 mM Tris pH 7.5, 100 mM KOAc pH 7.5, 2.5 mM $Mg(OAc)_2$, 0.25 mM spermidine, and 2 mM DTT) and contained 1 mM ATP and GTP (or GDPNP), 2 mM Upf1, 4 mM eRF1, 6 mM eRF3, and 3 nM termination complex. For subunit dissociation assays, 2 mM Rli1 was used as a positive control for splitting, and 50 mM PTH was added to the reactions to quantify dissociated complexes (PTH cannot access the peptidyl-tRNA unless released from the ribosome). Time points in both assays were quenched using 10% formic acid and run on electrophoretic TLC (Millipore). TLC plates were developed using a Typhoon FLA 9500 Phosphorimager

system and quantified using ImageQuantTL (GE Healthcare Life Sciences). Time courses were fit to single exponential kinetics using Kaleidagraph (Synergy Software).

In vitro reconstituted translation elongation

Translation elongation reactions were performed in 1X Buffer E (20 mM Tris pH 7.5, 100 mM KOAc pH 7.6, 2.5 mM $Mg(OAc)_2$, 0.25 mM Spermidine, and 2 mM DTT). Limited amounts of 80S initiation complexes (3 nM) were mixed with purified eEF1A (1 μM), aa-tRNA (500 nM), eEF2 (1 μM), eEF3 (1 μM), ATP (3 mM) and GTP (2 mM) in the presence or absence of eIF5A (1 μM). Reactions were incubated at 26^{o}C and time points quenched with 250 mM KOH. Peptide formation was monitored by electrophoretic TLC (Millipore). TLC plates were equilibrated with pyridine acetate buffer (5 ml pyridine, 200 ml acetic acid in 1 liter, pH 2.8) before electrophoresis at 1200 V for 28 min. Plates were developed using a Typhoon FLA 9500 Phosphorimager system (GE Healthcare Life Sciences) and quantified using ImageQuantTL (GE Healthcare Life Sciences). Time courses were fit to single exponential kinetics using Kaleidagraph (Synergy Software).

FIGURE LEGENDS

Figure 1. Yeast purified Upf1 constructs stimulate translation termination

(A) Schematic of *in vitro* termination assay. Termination complexes containing a [^{35}S]-Met-Phe-tRNA^{Phe} in the P site and a UAA codon in the A site are mixed with release factors and Upf1. Hydrolysis of the peptide by eRF1 releases a Met-Phe (MF) dipeptide that is analyzed via electrophoretic TLC and rate of release determined (example shown). (B) Rates of peptide release by eRF1:eRF3 and stimulation by Upf1. k_{obs} values are the means ± SEM (n=3). (C) Stimulation of subunit dissociation by Rli1 and Upf1 in PTH-coupled ribosome dissociation assay. (D) Coomassie stained gel of Upf1 constructs. (E) Termination rates for all Upf1 constructs. k_{obs} values represent the mean ± SEM (n=3).

Figure 2. Yeast specific contaminant of MBP purification stimulates translation termination

Peptidyl-hydrolysis assay comparing His-MBP purified from both *E. coli* and *S. cerevisiae* reveals presence of contaminating activity in yeast-purified sample. Blue box indicates normal rate of peptide release while red box indicates a stimulated reaction.

Figure 3. Contaminating activity is caused by a heat-insensitive protein

Peptidyl-hydrolysis assay comparing *S. cerevisiae* purified Upf1 CH domain before and after heat treatment, as well as in the presence of proteinase K. Blue box indicates normal rate of peptide release while red box indicates a stimulated reaction.

Figure 4. Purified eIF5A can stimulate translation termination and is heat-insensitive

Peptidyl-hydrolysis assay in the presence of various amounts of hyspusinated eIF5A, hypusinated eIF5A that was boiled at 95^{o}C for 15 min prior to addition, unmodified eIF5A, and a purified His_6-MBP construct expressed in *S. cerevisiae* containing low-level (though highly active) eIF5A contamination. Blue box indicates normal rate of peptide release while red box indicates a stimulated reaction.

Figure 5. eIF5A contaminates eEF2 fractions off S200 column

Elongation kinetics of MPPK formation using different fractions of eEF2 after gel filtration chromatography. Early fractions (#1-2) contain little to no contaminant eIF5A while later fractions (#3-4) have some level of eIF5A contamination.

REFERENCES

Amrani, N., Dong, S., He, F., Ganesan, R., Ghosh, S., Kervestin, S., Li, C., Mangus, D.A., Spatrick, P., and Jacobson, A. (2006). Aberrant termination triggers nonsense-mediated mRNA decay. Biochemical Society transactions *34*, 39-42.

Celik, A., Kervestin, S., and Jacobson, A. (2015). NMD: At the crossroads between translation termination and ribosome recycling. Biochimie *114*, 2-9.

Eyler, D.E., and Green, R. (2011). Distinct response of yeast ribosomes to a miscoding event during translation. RNA *17*, 925-932.

Eyler, D.E., Wehner, K.A., and Green, R. (2013). Eukaryotic release factor 3 is required for multiple turnovers of peptide release catalysis by eukaryotic release factor 1. The Journal of biological chemistry *288*, 29530-29538.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Molecular cell *51*, 35-45.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nature methods *11*, 319-324.

Park, M.H., Cooper, H.L., and Folk, J.E. (1981). Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. Proceedings of the National Academy of Sciences of the United States of America *78*, 2869-2873.

Schmidt, C., Becker, T., Heuer, A., Braunger, K., Shanmuganathan, V., Pech, M., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2016). Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. Nucleic acids research *44*, 1944-1951.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proceedings of the National Academy of Sciences of the United States of America *108*, E1392-1398.

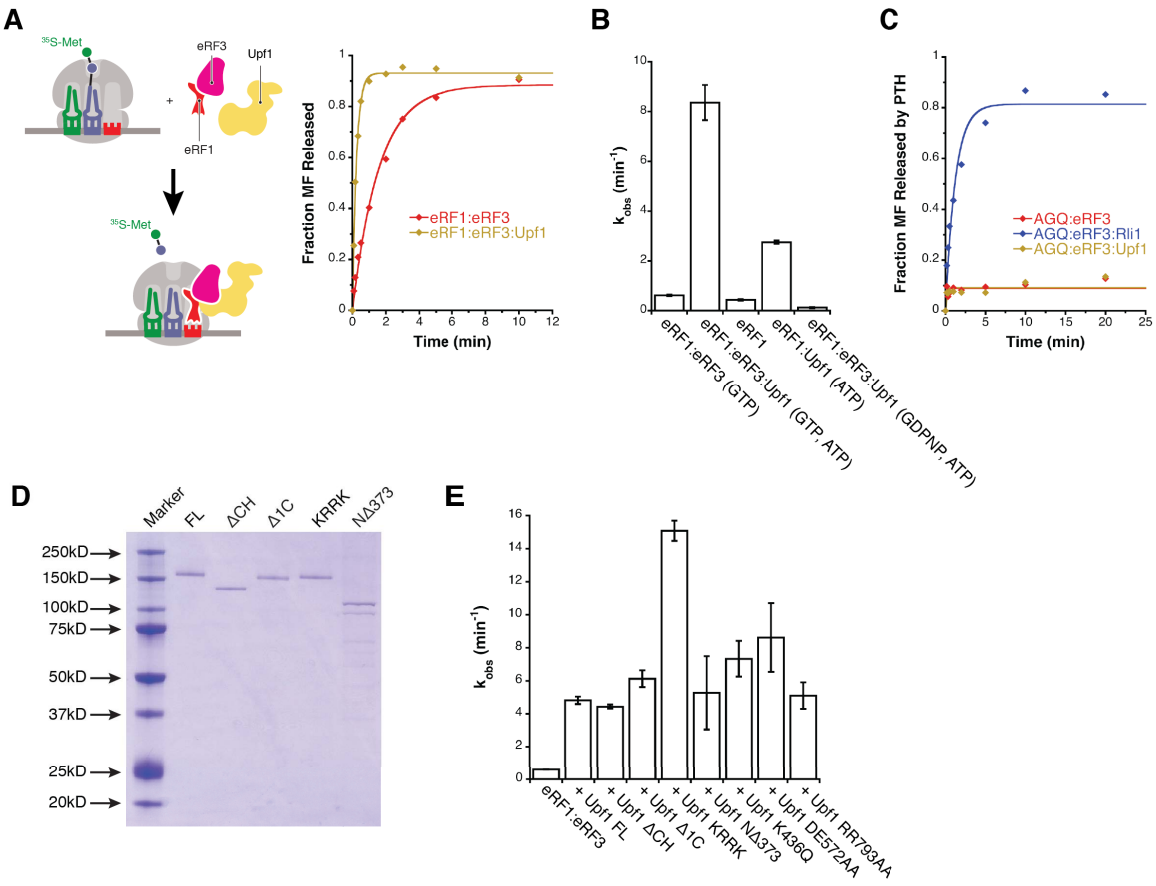

A
35S-Met
eRF3
Upf1
eRF1
35S-Met
Fraction MF Released
Time (min)
eRF1:eRF3
eRF1:eRF3:Upf1
B
kobs (min-1)
eRF1:eRF3 (GTP)
eRF1:eRF3:Upf1 (GTP, ATP)
eRF1
eRF1:Upf1 (ATP)
eRF1:eRF3:Upf1 (GDPNP, ATP)
C
Fraction MF Released by PTH
Time (min)
AGQ:eRF3
AGQ:eRF3:Rli1
AGQ:eRF3:Upf1
D
Marker
FL
ΔCH
Δ1C
KRRK
NΔ373
250kD
150kD
100kD
75kD
50kD
37kD
25kD
20kD
E
kobs (min-1)
eRF1:eRF3
+ Upf1 FL
+ Upf1 ΔCH
+ Upf1 Δ1C
+ Upf1 KRRK
+ Upf1 NΔ373
+ Upf1 K436Q
+ Upf1 DE572AA
+ Upf1 RR793AA

Figure 1

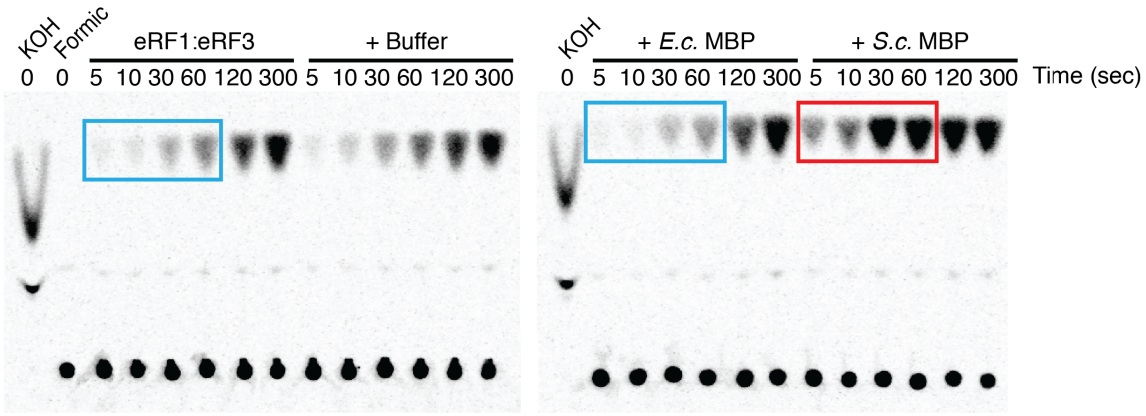

KOH
Formic
eRF1:eRF3
+ Buffer
0 0 5 10 30 60 120 300 5 10 30 60 120 300
KOH
+ E.c. MBP
+ S.c. MBP
0 5 10 30 60 120 300 5 10 30 60 120 300
Time (sec)

Figure 2

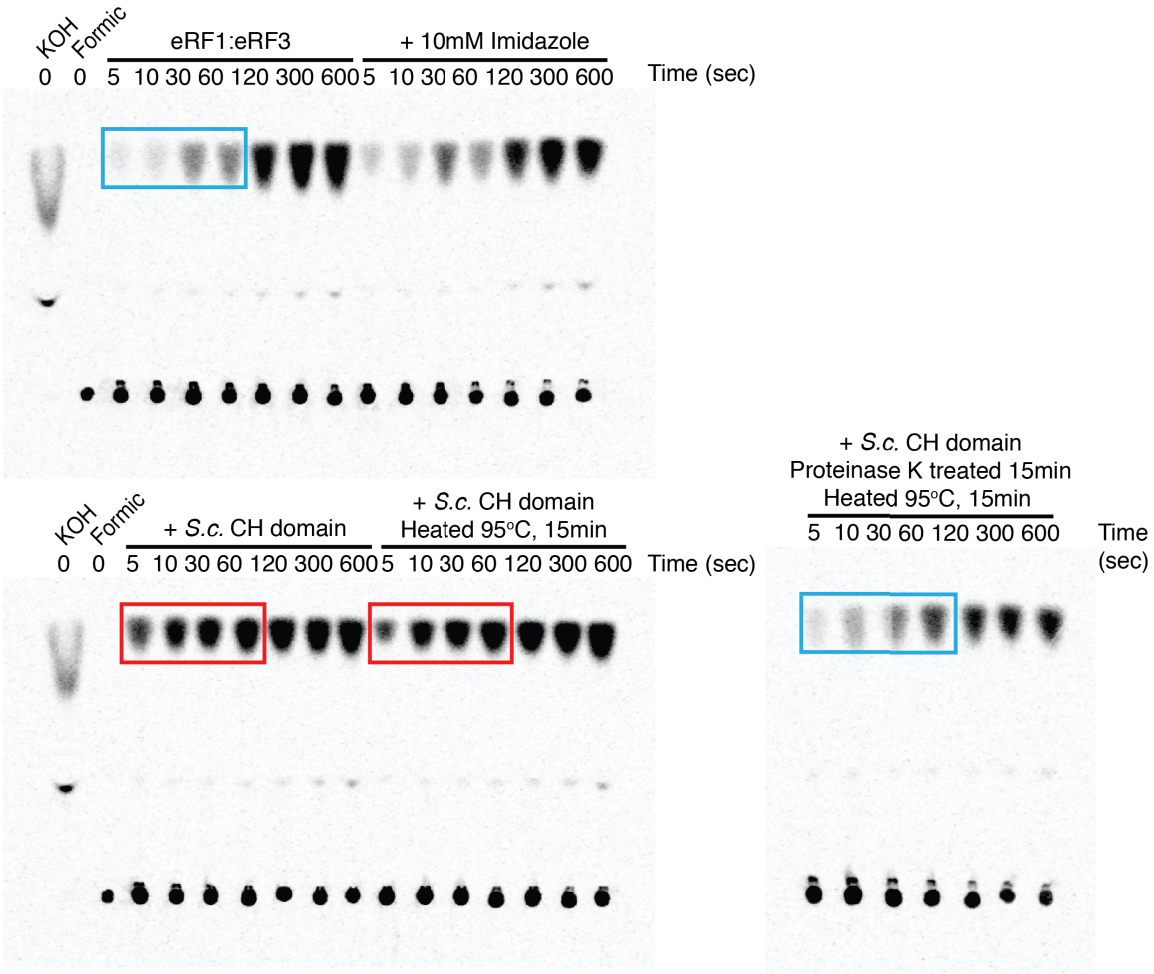

KOH
Formic
eRF1:eRF3
+ 10mM Imidazole
0 0 5 10 30 60 120 300 600 5 10 30 60 120 300 600
Time (sec)
KOH
Formic
+ S.c. CH domain
+ S.c. CH domain
Heated 95°C, 15min
0 0 5 10 30 60 120 300 600 5 10 30 60 120 300 600
Time (sec)
+ S.c. CH domain
Proteinase K treated 15min
Heated 95°C, 15min
5 10 30 60 120 300 600
Time
(sec)

Figure 3

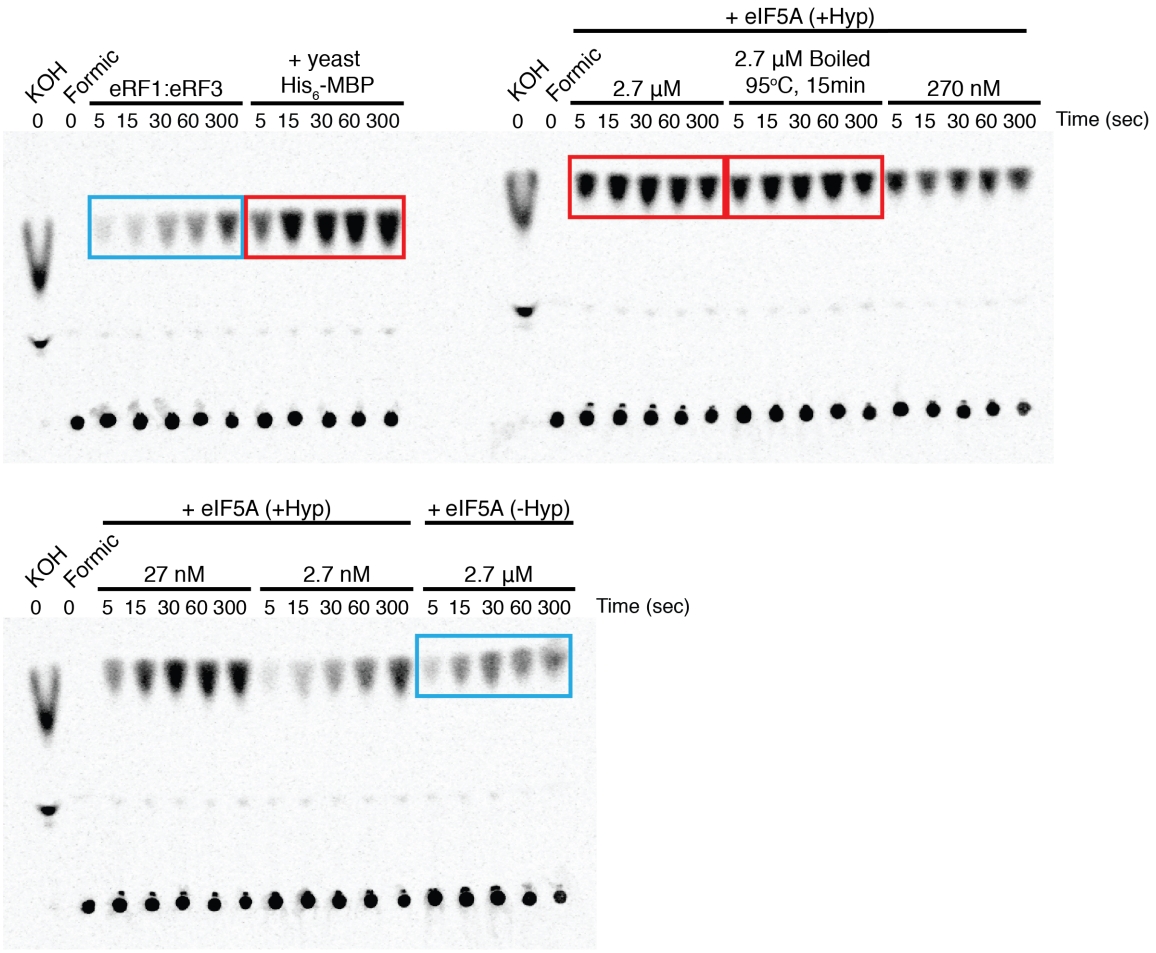

+ eIF5A (+Hyp)
KOH
Formic
eRF1:eRF3
+ yeast
His6-MBP
2.7 µM
2.7 µM Boiled
95°C, 15min
270 nM
0 0 5 15 30 60 300 5 15 30 60 300
0 0 5 15 30 60 300 5 15 30 60 300 5 15 30 60 300
Time (sec)
+ eIF5A (+Hyp)
+ eIF5A (-Hyp)
KOH
Formic
27 nM
2.7 nM
2.7 µM
0 0 5 15 30 60 300 5 15 30 60 300 5 15 30 60 300
Time (sec)

Figure 4

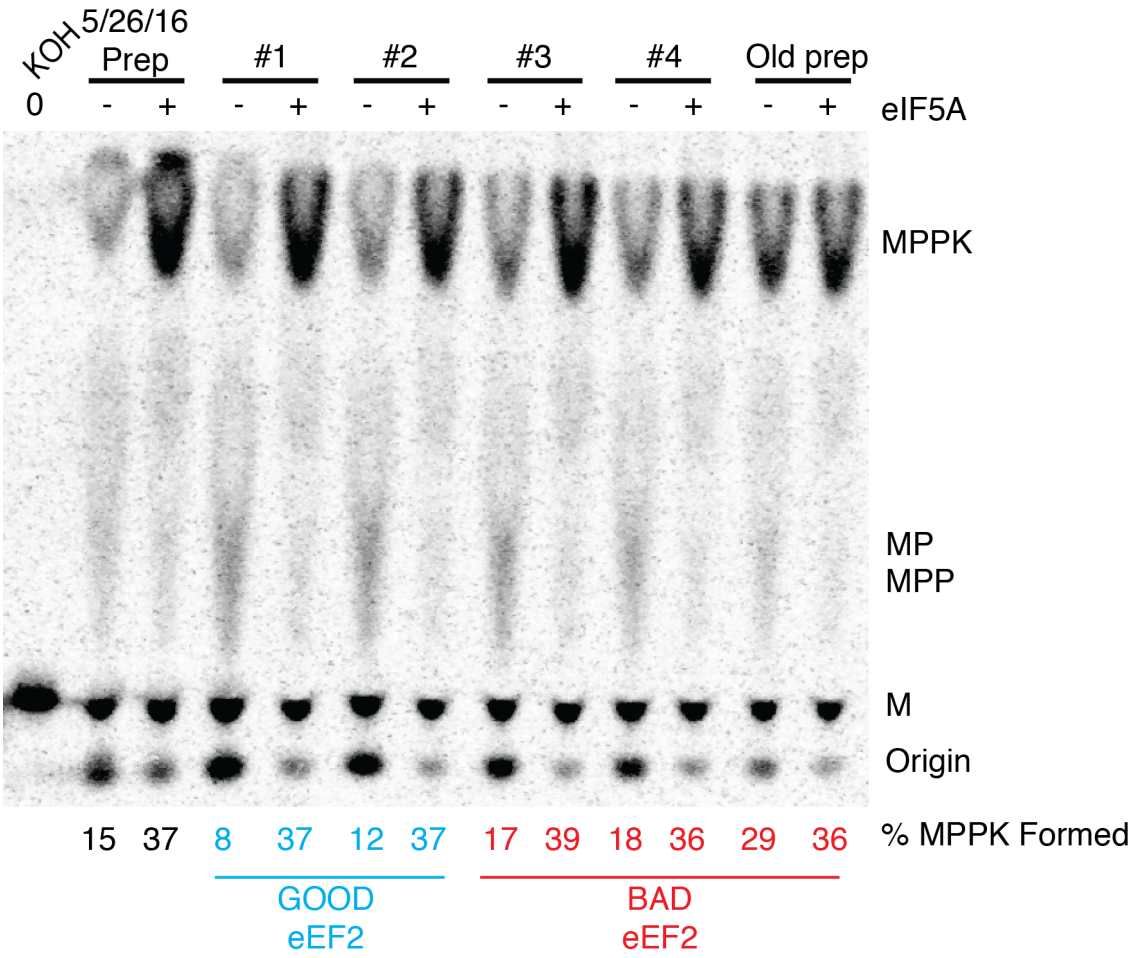

KOH
5/26/16 Prep
#1
#2
#3
#4
Old prep
0
- + - + - + - + - + - +
eIF5A
MPPK
MP
MPP
M
Origin
15 37 8 37 12 37 17 39 18 36 29 36
% MPPK Formed
GOOD eEF2
BAD eEF2

Figure 5

Curriculum Vitae

Anthony P. Schuller

aschull2@jhmi.edu

Academic Education

2017 Ph.D., Program in Biochemistry, Cellular and Molecular Biology
Department of Molecular Biology and Genetics
Johns Hopkins University School of Medicine, Baltimore, MD

2011 B.A., Biochemistry
Department of Chemistry
University of Pennsylvania, Philadelphia, PA

Research Experience

2017-2011 Ph.D. Dissertation Research
Department of Molecular Biology and Genetics
Johns Hopkins University School of Medicine, Baltimore, MD
Advisor: Dr. Rachel Green
Topic: Eukaryotic translation and quality control

2007-2011 Undergraduate Research
Gene Expression and Regulation Program
The Wistar Institute, Philadelphia, PA
Advisor: Dr. Emmanuel Skordalakes
Topic: Telomerase Structure and Function

Awards/Honors/Fellowships

2013 NSF Graduate Research Fellowship Honorable Mention
2012 NSF Graduate Research Fellowship Honorable Mention
2011 University Scholar (University of Pennsylvania)

Publications

1. **Schuller, A.P.** Zinshteyn, B., Enam, S.U. & Green, R. Directed hydroxyl radical probing reveals Upf1 binding to the 80S ribosomal E site rRNA at the L1 stalk. *In review.*
2. **Schuller, A.P.** & Green, R. The ABC(E1)s of ribosome recycling and reinitiation. *Molecular Cell* 66, 578-580 (2017).

3. McClary, B., Zinshteyn, B., Meyer, M., Jouanneau, M., Pellegrino, S., Yusupova, G., **Schuller, A.**, Reyes, J.C.P., Lu, J., Guo, Z., Ayinde, S., Luo, C., Dang, Y., Romo, D., Yusupov, M., Green, R. & Liu, J.O. Inhibition of eukaryotic translation by the antitumor natural product Agelastatin A. *Cell Chemical Biology* 24, 605-613 (2017).
4. **Schuller, A.P.***, Wu, C.C.*, Dever, T.E., Buskirk, A.R. & Green, R. eIF5A functions globally in translation elongation and termination. *Molecular Cell* 66, 194-205 (2017).
    - Spotlight in *Trends in Biochemical Sciences* 42, 587-588 (2017).
5. Kowalinski, E., **Schuller, A.**, Green, R & Conti, E. *Saccharomyces cerevisiae* Ski7 is a GTP-binding protein adopting the characteristic conformation of active translational GTPases. *Structure* 23, 1-8 (2015). PMCID: PMC4509514.
6. Koutmou, K.S., **Schuller, A.P.**, Brunelle, J.L., Radhakrishnan, A., Djuranovic, S. & Green, R. Ribosomes slide on lysine-encoding homopolymeric A stretches. *eLife,* e05534 (2015). PMCID: PMC4363877.
7. Van den Elzen, A.M.G, **Schuller, A.**, Green, R. & Séraphin, B. Dom34-Hbs1 mediated dissociation of inactive 80S ribosomes promotes restart of translation after stress. *EMBO J.* 33, 265-276 (2014). PMCID: PMC3989619.
8. **Schuller, A.P.**, Harkisheimer, M.J. & Skordalakes, E. *In vitro* reconstitution of the active *T. castaneum* telomerase. *J. Vis. Exp*. 53, e2799, doi: 10.3791/2799 (2011). PMCID: PMC3196162.
9. Mason, M., **Schuller, A.** & Skordalakes, E. Telomerase structure function. *Curr. Opin. Struct. Biol.* 21, 92-100 (2011).
10. Gillis, A.J., **Schuller, A.P.** & Skordalakes, E. Structure of the *Tribolium castaneum* telomerase catalytic subunit TERT. *Nature* 455, 633–637 (2008).

Presentations/Abstracts

1. **Schuller, A.**, Wu, C., & Green, R. eIF5A functions globally in translation elongation and termination. NCI Symposium on RNA Biology, Bethesda, MD (2017). Poster presentation.
2. **Schuller, A**., Wu, C., Dever, T. & Green, R. eIF5A functions as both an elongation and termination factor in yeast. Translational Control Meeting, Cold Spring Harbor, NY (2016). Oral presentation.
3. **Schuller, A**., Wu, C., Guydosh, N. & Green, R. Upf1 Interacts with Release Factors to Stimulate Peptide Hydrolysis at Termination Codons. EMBO Protein Synthesis and Translational Control, Heidelberg, Germany (2015). Oral presentation.
4. **Schuller, A.** & Green, R. Functional analysis of Upf1 with the Ribosome. HHMI Meeting, Chevy Chase, MD (2015). Poster presentation.

5. **Schuller, A.** & Green, R. Functional analysis of Upf1:Upf2 with the Ribosome. Translational Control Meeting, Cold Spring Harbor, NY (2014). Poster presentation.
6. **Schuller, A.** & Green, R. Upf1 Promotes Ribosomal Subunit Dissociation *in vitro*. Ribosomes Conference, Napa, CA (2013). Poster presentation.
7. Koutmou, K.S., **Schuller, A.** & Green, R. Re-exploring mechanistic connections between mRNA surveillance pathways. Translational Control Meeting, Cold Spring Harbor, NY (2012). Poster presentation.

Teaching Experience

2014-2012 Molecular Biology and Genomics Graduate Teaching Assistant

2014 Lecture in Molecular Biology and Genomics course
Topic: Methods in Molecular Biology